UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA

Evaluation of Land Surface Temperature in Atmospheric Reanalyses using Earth Observations

João Frederico Gonçalves Johannsen

Mestrado em Ciências Geofísicas
Especialização em Meteorologia e Oceanografia

Dissertação orientada por:
Emanuel Dutra, Sofia Ermida

2019

Acknowledgements

Thanks are due to my thesis advisors, Emanuel and Sofia, for their precious assistance throughout the entire process that culminated in this work. I would also like to thank my family and friends for their unconditional support during my master's degree. Finally, thanks are due to my English teacher Paul, whose writing advices helped correcting the writing and made it more readable.

This work was funded by FCT under project CONTROL: PTDC/CTA-MET/28946/2017.

Abstract:

The surface skin temperature (SKT) is a key variable in surface-atmosphere energy and water exchanges. In the first part of this work, the SKT from two reanalyses (ERA-Interim and ERA5) of the European Centre for Medium-Range Weather Forecasts (ECMWF) is evaluated against satellite-based Land Surface Temperature (LST) retrieved by the Satellite Application Facility on Land Surface Analysis (LSA-SAF), during the 2004-2015 period over the Iberian Peninsula. Offline simulations by the Hydrology Tiled ECMWF Scheme of Surface Exchanges over Land (HTESSEL) model are also assessed. We apply four performance metrics to the daily maximum and minimum SKT: the mean error, standard deviation of the error, temporal correlation and root mean squared error. The results show an underestimation of the daytime SKT and a small overestimation of nighttime SKT in all the products, which is in line with previous studies. In general, ERA5 presents a consistent improvement over ERA-Interim by showing an overall better agreement with the satellite observations. There is also a reasonably high correlation between the misrepresentation of vegetation cover in the HTESSEL model and the daily maximum SKT bias. In the second part of the thesis, we apply changes to some parameters of the HTESSEL model. The parameters' impact in the simulation of SKT is then assessed by comparing them to the satellite-LST in a confined domain centred in Évora. The implementation of a revised model representation of vegetation cover (based on the ESA-CCI Land Cover Dataset) is shown to have a positive impact on SKT, especially during daytime. A new soil discretization scheme, on the other hand, does not significantly impact the simulation of SKT. Finally, a sensitivity study applied to the c_{veg} parameter (the model vegetation cover) reaffirms the importance of the representation of vegetation in the model, as there is a correlation between c_{veg} and the daily maximum SKT bias with the revised model vegetation (while the same correlation cannot be reproduced with the original model vegetation).

KEY WORDS: Surface skin temperature, ECMWF reanalyses, LSA-SAF's LST, vegetation cover

Resumo:

A temperatura da superfície terrestre (SKT) é um dos parâmetros-chave nas trocas terra-atmosfera de energia e água. Além disso, é cada vez mais importante na assimilação de dados e na parametrização de variáveis em modelos. Estudos anteriores apontam para subestimações consideráveis da SKT em reanálises, principalmente durante o dia e em regiões semi-áridas e áridas.

Na primeira parte desta tese, a SKT de duas reanálises (ERA-Interim e ERA5) do Centro Europeu de Previsão de Tempo a Médio-Prazo (ECMWF) é avaliada, tendo como produto de referência a temperatura da superfície terrestre (LST) obtida por satélite pelos Serviços de Aplicações de Satélite a Análises da Superfície Terrestre (LSA-SAF). De igual forma, é avaliada a SKT de simulações *offline* de uma versão do modelo HTESSEL do ECMWF muito semelhante à utilizada na concepção da ERA5. Este estudo é realizado entre 2004 e 2015 (período em que a LST reprocessada se encontra disponível) sobre a Península Ibérica, numa resolução espacial de $0.25º \times 0.25º$. Como a LST de satélite tem uma resolução maior (~5 km), é realizado o *upscaling* dos dados de satélite para se poder fazer a comparação com as reanálises. Além disso, esta base de dados realiza medições num determinado pixel apenas quando não há nuvens sobre esse pixel. Por causa disso, aplica-se um limite máximo de cobertura nebulosa de 0.3 em cada pixel e para cada produto, de forma a evitar uma contaminação elevada dos dados por nuvens e, ao mesmo tempo, manter uma percentagem considerável de dados válidos. Quatro métricas de desempenho são aplicadas à SKT máxima e mínima diária: erro médio, desvio-padrão do erro, correlação temporal e a raiz quadrada do erro médio quadrático. Para facilitar a organização dos resultados, aplica-se o algoritmo K-Means de forma a dividir o domínio em diferentes clusters. Os clusters são determinados a partir do ciclo diurno médio da LST no Verão dos doze anos considerados no estudo. Os resultados mostram, em todos os produtos, uma subestimação considerável da SKT de dia

e uma sobrestimação pouco significativa de noite, em linha com estudos anteriores, sendo essa disparidade superior em regiões mais áridas. No geral, a ERA5 apresenta um produto com qualidade superior em relação à ERA-Interim, pois é a reanálise que mais se aproxima das observações de satélite. Mostra-se também que existe uma correlação razoavelmente elevada entre o erro na representação da cobertura vegetal no modelo (ao comparar a fracção de cobertura vegetal do modelo pela observada pela base de dados da Copernicus) e o erro médio na simulação da SKT.

Na segunda parte deste trabalho, são aplicadas mudanças a alguns parâmetros do modelo HTESSEL e o seu impacto na simulação da SKT é avaliado pelas observações de satélite da LST, num domínio mais restrito centrado em Évora (quatro pontos na resolução original de 0.25º × 0.25º). A implementação de uma representação mais realista de coberto vegetal no modelo é obtida a partir da base de dados de cobertura terrestre da ESA-CCI, mantendo os tipos de vegetação alta/baixa originais considerados pelo modelo mas alterando a respectiva fracção de cobertura vegetal. O domínio passa a ser coberto maioritariamente por vegetação baixa, ao contrário do domínio original que era coberto quase na totalidade por vegetação alta. Esta nova representação da vegetação tem um impacto positivo na SKT durante o dia, e uma nova discretização do solo (nove camadas em vez de quatro) reduz ainda mais o erro, embora com um impacto menos acentuado do que a introdução da nova vegetação. Ainda assim, mesmo depois destas alterações, o viés nocturno e a diferença de fase durante o dia mantêm-se no ciclo diurno médio da SKT. É importante referir que o tipo de vegetação baixa considerado pelo modelo HTESSEL não é directamente equivalente ao obtido com os dados da ESA-CCI e ao utilizar outros tipos de vegetação baixa no modelo, estes originam resultados menos positivos. Isto dever-se-á ao facto da cobertura vegetal ser menor no tipo de vegetação baixa originalmente considerado pelo modelo quando comparado com os restantes tipos existentes, o que implica um aumento da fracção de cobertura vegetal e, consequentemente, um aumento do viés da SKT quando se muda para outro tipo de vegetação baixa.

Finalmente, um estudo de sensibilidade é aplicado ao pârametro c_{veg}, que representa a cobertura vegetal do modelo, de forma a verificar a sua relação com o viés na simulação da SKT, utilizando a representação de vegetação original do modelo HTESSEL e, seguidamente, a representação revista. O parâmetro c_{veg} é perturbado com valores entre 0.1 e 1, para os tipos de vegetação alta e baixa, formando no total 100 pares de c_{veg}. Esses pares são obtidos através de uma distribuição quasi-aleatória, distribuição de Sobol, que permite preencher na totalidade o domínio considerado, sem introduzir correlações entre as diferentes perturbações e sem criar clusters e espaços vazios no domínio. Este estudo reafirma a importância da representação da vegetação no modelo, pois existe uma correlação entre c_{veg} e o viés da SKT máxima diária quando se considera a vegetação revista (ao contrário do que acontece com a vegetação original do modelo, em que não se verifica qualquer correlação com o erro).

As simulações desacopladas (offline) permitem avaliar o impacto dos parâmetros de superfície na simulação de SKT, mas é igualmente importante estudar o efeito da vegetação revista em simulação acopladas com a atmosfera. Também é importante referir que as mudanças na cobertura de vegetação afectam o balanço de água (que não é estudado neste trabalho) e causam outras alterações noutras estações do ano. Contudo, a pouca disponibilidade de dados de LST de satélite e de outras observações (como, por exemplo, fluxos e temperatura e humidade do solo) dificultam eventuais estudos adicionais.

Ainda que a ERA5 demonstre ser um produto com melhor qualidade que a ERA-Interim, os resultados sugerem a necessidade de uma revisão da vegetação no modelo HTESSEL sobre a Península Ibérica, nomeadamente as fracções de cobertura de vegetação baixa / alta. Do mesmo modo, a própria definição dos tipos de vegetação baixa / alta e o valor de c_{veg} associado poderão também precisar de uma revisão.

PALAVRAS-CHAVE: Temperatura da superfície terrestre, reanálises do ECMWF, LST da LSA-SAF, cobertura vegetal

Table of Contents

List of Tables

List of Figures

1 - Introduction

The Land Surface Temperature (LST) is a key variable in surface-atmosphere energy exchange processes and it was recently integrated as an Essential Climate Variable (ECV) into the Global Climate Observing System (GCOS). ECVs are physical, chemical, and biological variables that are considered crucial in the characterization of the global climate (Bojinski et al., 2014).

LST observations are obtained from measurements of the radiance emitted by the land surface, usually performed in the thermal infrared (TIR) or in the microwave (MW) regions of the electromagnetic spectrum, by sensors orbiting the Earth on satellites or in in-situ meteorological stations.

The LST obtained in the TIR channel is defined as the radiometric temperature (due to its derivation from the radiance emitted by the planet's surface), which is the temperature of the surface layer equal to the penetration depth of the radiation used in its calculation (Norman and Becker, 1995; Li et al., 2013). In the case of TIR radiation, the penetration depth is a few millimetres (Wan, 1999).

There are several methods to derive LST from remote sensing observations. In the TIR region, LST can be obtained using many different types of methods (e.g. the "single channel" approach, the "split-window" algorithm; Li et al., 2013). Regarding MW methods, these are split in the following three categories: empirical methods, physical-based methods, and artificial neural network methods (Han et al., 2019). Products based on TIR algorithms have higher resolution and overall higher quality than MW-derived products (e.g. Prigent et al., 2016). One of the limitations of TIR-based LST data (and in relation to MW-based LST) is its dependence on clear-sky weather conditions. Absence of LST data will occur over a given pixel if the sky is cloudy during the period in which the satellite is over that pixel.

In this study, we use LST produced by the Satellite Application Facility on Land Surface Analysis (LSA-SAF). This observational dataset is derived from the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) onboard the Meteosat Second Generation (MSG) series of geostationary satellites. The observations are retrieved, under clear-sky conditions, every 15 minutes for the full MSG disk that covers Africa, Europe and part of South America.

Good (2016) showed that on cloudy and/or windy days and on clear sky nights, the LST and the two-metre air temperature (T2m) have very similar values, whereas the largest contrasts happen during cloud-free days. Therefore, for certain weather conditions, LST can serve as a predictor of the T2m, a widely used variable in climate change studies and a key piece of evidence of global warming. This is quite relevant due to LST's globally available datasets, while T2m is only measured in in-situ stations.

Land surface is one of the main components of the earth's climate system. Its interaction with the atmosphere involves energy fluxes and water and carbon exchanges that are crucial for weather forecasting and climate studies (e.g. Schär et al., 2004; Seneviratne et al., 2006). In spite of their importance, the land-atmosphere exchanges in land surface models present considerable biases, especially during extreme weather events (Ukkola et al., 2016; Kala et al., 2016). In the case of simulating latent and sensible heat fluxes, the physics-based land surface models were outperformed by simplistic empirical models (Best et al., 2015).

Many climate reanalyses include a variable denominated SKin Temperature (SKT) which is essentially the model equivalent to LST. A climate reanalysis is a combination of model output and observations. It contains estimates of numerous atmospheric and surface variables (e.g. air temperature, pressure, rainfall) and it is a global product available for long periods of time, from a few decades to over a century

of data. It gives a numerical description of the recent climate, allowing scientists to conduct studies on a variety of climate-related subjects (Dee et al., 2011).

The European Centre for Medium-Range Weather Forecasts (ECMWF) has developed many atmospheric and ocean reanalyses throughout the years (starting in 1979), with the two most recent atmospheric reanalyses being ERA-Interim and ERA5 (Hersbach et al., 2018). The skill of these two reanalyses in terms of simulating SKT is analysed in this study.

Since ERA5 is a fairly recent product, it has not yet been featured widely in scientific literature. Preliminary results conducted by ECMWF showed ERA5's overall improvement in comparison to ERA-Interim in simulating several different variables (Hersbach et al., 2018). Albergel et al. (2018) compared ERA5 and ERA-Interim atmospheric forcing in land surface model simulations, ultimately showing that ERA5 provides an improved product over ERA-Interim. Besides being recent, ERA5 will serve as the official replacement of ERA-Interim (by the end of 2019), hence it is imperative to evaluate its ability in simulating an ECV like the SKT.

The ECMWF reanalyses are generated by the Integrated Forecasting System (IFS), a global data assimilation and forecasting system developed by ECMWF. The Hydrology Tiled ECMWF Scheme of Surface Exchanges over Land (HTESSEL) is one component of the IFS, representing the surface skin layer — a shallow layer with zero heat capacity that separates the subsoil from the atmosphere and intercepts and emits radiation. Over the years, the HTESSEL model has been updated with newer implementations, including a new snow scheme, vegetation seasonality and a revised soil hydrology (Balsamo et al., 2011).

Evaluation of simulated SKT with LST observations has been the subject of analysis in many previous articles (e.g. Trigo et al., 2015; Zhou et al., 2017; Orth et al., 2017). Trigo et al. (2015) found an underestimation of daytime SKT over most of Africa and Europe (especially over semiarid regions) and a slight SKT overestimation during nighttime for the ECMWF HTESSEL model (forced with ERA-Interim) when compared to LSA-SAF's LST. With the same datasets, focusing on Europe but extending the temporal range, Orth et al. (2017) also found an underestimation of the SKT daily range (especially in the Iberian Peninsula, < -10ºC). Zhou et al. (2017) examined several reanalysis products (including ERA-Interim) over China with the reference-LST measured by in-situ stations. Similar to the previous studies, most of the reanalyses showed an underestimation of the SKT, especially in the summer and in arid regions.

LST can also help constrain surface variables. Trigo et al. (2015) presented a revision for different surface parameters (Leaf Area Index (LAI), roughness length for momentum and for heat) and assessed its impact in the simulation of SKT. The revised roughness lengths had a positive impact on the daytime SKT while the revised LAI had a minor yet positive effect. Orth et al. (2017) showed that the SKT performance is highly sensitive to three surface parameters: the minimum stomatal resistance, the skin conductivity, and the soil moisture stress function.

Moreover, LST is pertinent in data assimilation, which is the process of minimizing model errors by updating the input variables at each time step if observations of those variables are available. For example, Ghent et al. (2010) showed that the assimilation of satellite-LST positively impacts the simulation of SKT, soil moisture and the latent and sensible heat fluxes.

The first objective of this thesis is to evaluate different ECMWF products, namely the reanalyses ERA-Interim and ERA5, and the HTESSEL model, in terms of simulated SKT in relation to the satellite-LST over the Iberian Peninsula. The second objective is to improve the simulation of SKT of the HTESSEL model by running offline (uncoupled) simulations in a confined domain centred in Évora. Our

hypothesis is that certain model parameters (e.g. the prescribed land cover) can be improved, and that their enhancement leads to a positive impact in the simulation of SKT. The following chapter presents the data and methods applied in this work, followed by the results and discussion. The last chapter presents the final conclusions of this thesis.

2 - Data and Methods

2.1 - Data:

2.1.1 - ECMWF's HTESSEL surface model

The ECMWF land surface model HTESSEL is one of the components of the ECMWF IFS. HTESSEL represents the surface skin layer, a shallow layer with zero heat capacity that separates the subsoil from the atmosphere and intercepts and emits radiation. Each grid point of this layer can be divided into different tiles that represent different types of land cover (bare ground, low and high vegetation, intercepted water (on the canopy), and shaded and exposed snow). Only the dominant type of low/high vegetation at each grid point is considered by the model. This information is then used to generate spatial fields of various parameters used in different parameterizations which are assumed to be dependent only on vegetation type (Table 2.1). A detailed description of the model assumptions and parameterization can be found online[1]. In the following, a more detailed description of the processes directly linked with the simulation of SKT in HTESSEL is presented. In this study, the version used was CY45R1, which is very close to the version used in ERA5 in terms of surface processes.

HTESSEL takes as input static 2-dimensional fields of low vegetation cover (CVL), high vegetation cover (CVH), dominant type of low vegetation (TVL) and dominant type of high vegetation (TVH). Neglecting interception and snow, the low vegetation tile fraction (c_{low}) and high vegetation tile fraction (c_{high}) are given by:

$$\begin{aligned} c_{low} &= CVL \times c_{veg}(TVL) \\ c_{high} &= CVH \times c_{veg}(TVH) \end{aligned} \tag{2.1}$$

Where c_{veg} is the vegetation cover which is dependent on vegetation type and given in Table 2.1. The total vegetation cover of a grid cell (HTESSEL-FCover) and bare ground (c_{bare}) are given by:

$$FCover = c_{low} + c_{high} \tag{2.2}$$

$$c_{bare} = 1 - FCover \tag{2.3}$$

The input data used in HTESSEL is the same as used in the reanalysis and operational weather forecasts and was derived from the Global Land Cover Characteristics (GLCC) data (Loveland et al., 2000).

[1] ECMWF: Part IV: Physical processes, ECMWF [online] Available from: https://www.ecmwf.int/en/elibrary/18714-part-iv-physical-processes (Accessed 16 May 2019).

Table 2.1. HTESSEL vegetation types and associated parameters' values. H/L differentiates low (L) from high (H) vegetation. $r_{s,min}$ is a minimum canopy resistance and z_{0m} and z_{0h} are the roughness lengths for momentum and heat, respectively (adapted from the HTESSEL Documentation).

Index	Vegetation type	H/L	$r_{s,min}$ (sm^{-1})	c_{veg}	z_{0m}	z_{0h}
1	Crops, mixed farming	L	100	0.90	0.25	$0.25\ 10^{-2}$
2	Short grass	L	100	0.85	0.20	$0.20\ 10^{-2}$
3	Evergreen needleleaf trees	H	250	0.90	2.0	2.0
4	Deciduous needleleaf trees	H	250	0.90	2.0	2.0
5	Deciduous broadleaf trees	H	175	0.90	2.0	2.0
6	Evergreen broadleaf trees	H	240	0.99	2.0	2.0
7	Tall grass	L	100	0.70	0.47	$0.47\ 10^{-2}$
8	Desert	—	250	0	0.013	$0.013\ 10^{-2}$
9	Tundra	L	80	0.50	0.034	$0.034\ 10^{-2}$
10	Irrigated crops	L	180	0.90	0.50	$0.50\ 10^{-2}$
11	Semidesert	L	150	0.10	0.17	$0.17\ 10^{-2}$
12	Ice caps and glaciers	—	—	—	$1.3\ 10^{-3}$	$1.3\ 10^{-4}$
13	Bogs and marshes	L	240	0.60	0.83	$0.83\ 10^{-2}$
14	Inland water	—	—	—	—	—
15	Ocean	—	—	—	—	—
16	Evergreen shrubs	L	225	0.50	0.10	$0.10\ 10^{-2}$
17	Deciduous shrubs	L	225	0.50	0.25	$0.25\ 10^{-2}$
18	Mixed forest/woodland	H	250	0.90	2.0	2.0
19	Interrupted forest	H	175	0.90	1.1	1.1
20	Water and land mixtures	L	150	0.60	—	—

The SKT is the temperature of the skin layer and it is computed from the surface energy balance equation calculated independently for each tile (represented by the letter *i*) following:

$$\left(1 - f_{Rs,i}\right)(1 - \alpha_i)R_S + \varepsilon\left(R_T - \sigma T_{sk,i}^4\right) + H_i + L_{v,s}E_i = \Lambda_{sk,i}\left(T_{sk,i} - T_1\right) \tag{2.4}$$

Where R_s and R_T are the downward shortwave radiation and longwave radiation, respectively, σ is the Stefan-Boltzmann constant, T_1 is the temperature of the upper soil (or snow) layer, $\Lambda_{sk,i}$ is the skin conductivity, $f_{Rs,i}$ is the fraction of net shortwave radiation transmitted directly to the top soil layer, α_i is the surface albedo and ε is the surface emissivity. The skin temperature is represented as $T_{sk,i}$. H_i is the sensible heat flux computed as:

$$H_i = \frac{\rho_a c_p}{r_a}\left(T_L + g z_L / c_p - T_{sk,i}\right) \tag{2.5}$$

Where ρ_a is the density of air, c_p is the specific heat of air, T_L and z_L are, respectively, the temperature and the height at the lowest atmospheric level of the model, and g is the acceleration of gravity. The aerodynamic resistance is given by $r_a = \left(|U_L| C_{H,i}\right)^{-1}$, where U_L is the horizontal wind speed at the lowest atmospheric level of the model and $C_{H,i}$ is the transfer coefficient for heat.

$L_{v,s}E_i$ is the latent heat flux from the skin layer, where $L_{v,s}$ is the latent heat of evaporation and E_i is the turbulent flux of water vapour given by:

$$E_i = \frac{\rho_a}{r_a + r_c}\left[q_L - q_{sat}(T_{sk,i})\right] \quad (2.6)$$

Where r_c is the canopy resistance, q_L is the specific humidity at the lowest atmospheric level and q_{sat} is the saturated specific humidity at the temperature $T_{sk,i}$.

The grid-box SKT is defined as the weighted average of the SKT on each tile fraction. The skin layer is thermally coupled to the four-layer soil below through the conductivity parameter ($\Lambda_{sk,i}$) related to the dominant vegetation type. The skin layer is coupled to the lowest level of the atmosphere using the Monin-Obukhov similarity theory (Viterbo and Beljaars, 1995) and this coupling is represented by $C_{H,i}$ (a function of atmospheric stability) and the roughness lengths for momentum and heat (z_{0m} and z_{0h}, respectively).

2.1.2 - ECMWF's Reanalyses

In the first part of this study, we assessed the simulation of SKT of two ECMWF Re-Analysis (ERA): ERA-Interim and ERA5.

The ECMWF reanalyses are generated by the IFS, a global assimilation and forecasting system developed by ECMWF. The ECMWF's IFS assimilates observations from multiple sources and model outputs of numerous climate variables in each time step in order to obtain the best estimate of each variable. The combined result is then used as the initial conditions for the next model run. This way, the historical observational dataset, which has a heterogeneous spatial distribution, is "reanalysed" and "completed" with the model first guess (previous forecast), providing a spatially homogeneous and coherent product that covers the whole planet, the analysis (Dee et al., 2011).

ERA-Interim is an atmospheric reanalysis based on a 2006 version of the IFS (cycle 31r2), and it served as a starting point to create a replacement for the previous ECMWF reanalysis, ERA-40. Its configuration used a 30-minute time step and a spectral TL255 horizontal resolution (approximately 79 km on a reduced Gaussian grid). The vertical resolution has 60 model layers that reach the top of the atmosphere, located at 0.1 hPa. The surface fields have a three-hourly resolution (eight daily values; Dee et al., 2011).

ERA5 is the latest ECMWF's atmospheric reanalysis, produced by Copernicus Climate Change Service. It is based on a 2016 version of the IFS (cycle 41r2). The horizontal resolution is about 31 km (TL639). It has 137 vertical layers culminating at 0.01 hPa. The analysis and forecast fields have 24 daily values (hourly output). It is the official replacement of ERA-Interim, offering a global improvement with several different technical changes (Hersbach et al., 2019; see Table 2.2) and innovations. The main innovations included new reanalysed variables and an uncertainty estimate for the reanalysis results (Hersbach et al., 2018).

The SKT data was extracted from the ECMWF server and converted to a regular 0.25º × 0.25º grid. In addition to the SKT, we also extracted the Total Cloud Cover (TCC), which is a variable that quantifies the percentage of cloud cover in a given grid point.

Table 2.2. Main differences between ERA5 and ERA-Interim (adapted from Hersbach et al., 2019).

	ERA-Interim	ERA5
Period	1979 – until end of August 2019	1950 – present
IFS model cycle	31r2 (2006)	41r2 (2016)
Data assimilation	12-hour 4D-Var	12-hour 4D-Var ensemble
Horizontal resolution	79 km (TL255)	31 km (TL639)
Vertical resolution	60 levels to 0.1 hPa	137 levels to 0.01 hPa
Land-surface model	TESSEL	HTESSEL
Uncertainty estimate	none	From the 4D-Var ensemble, 10 members at 63 km (TL319)
Output frequency	6-hourly for analyses 3-hourly for forecasts	Hourly (uncertainty 3-hourly)

2.1.3 - Simulations setup

HTESSEL source code is available as an independent library from the atmospheric model (also referred to as "externalized"). HTESSEL's externalization allows it to perform land-surface only (or offline) simulations at a much lower computation cost when compared to full global atmospheric simulations. The offline simulations are forced by near-surface state of air temperature, humidity, wind speed, pressure, solar and thermal downwelling energy, and precipitation. This data can be from observations or taken from the lowest model level of atmospheric forecasts of reanalysis (e.g. Balsamo et al., 2015).

In the first part of the work, the initial conditions were obtained for both ERA-Interim and ERA-5 reanalyses considered in this study. The simulations started in 2002 to allow the model to spin up (that is, to reach equilibrium), running for a 16-year interval until the end of 2017. The simulations consisted of monthly chunks using the restart facility with a 15-minute time step. Initial simulations with a 1-hour time step, which is commonly used, indicated some temporal lag in the SKT simulations associated with the numerical solver.

In the second part of the thesis, we performed uncoupled simulations forced only with ERA5, starting in 2009 and ending in September 2010, for a selected number of grid-points.

2.1.4 - LSA-SAF's Land Surface Temperature

The LST disseminated by the LSA-SAF is derived from measurements performed by the SEVIRI onboard the MSG series of satellites by employing a generalized "split-window" technique. This method estimates LST as a linear function of the brightness temperatures at the top of the atmosphere measured by SEVIRI's IR channels centred at 10.8 μm and 12.0 μm. The regression coefficients depend explicitly on the surface emissivity for both channels and implicitly on the total column water vapor and the satellite zenith view angle (SZA) (Trigo et al., 2011).

The LST is available every 15 minutes for all the land pixels of the MSG disk (which comprises SZAs between 0º and 80º), with a resolution of 3 km at the sub-satellite (nadir) point. LST uncertainty is usually between 1–2 K, except for regions near the edge of the MSG disk (due to large optical paths associated with high SZAs) or arid areas (where the surface emissivity's uncertainty is generally high, e.g. the Sahara desert) in which the error is larger (Trigo et al., 2011).

We used the LST obtained by remote sensing to evaluate the quality of both ERA products. The LST data was extracted from the LSA-SAF server located in the Portuguese Weather Institute's (IPMA)

headquarters. For the comparison between LST and SKT to be consistent, we performed an upscaling of the LST data, by computing the median of the whole group of LST pixels within each 0.25º × 0.25º grid cell. The number of original LST data (~ 5 km of resolution) in each grid cell varied between 30 and 56 pixels.

2.1.5 - Land Cover Datasets

Two different datasets were used in this study: the Copernicus Global Land Service (CGLS) fraction of vegetation cover (CGLS-FCover) and the European Space Agency's Climate Change Initiative (ESA-CCI) Land Cover dataset.

The CGLS-FCover represents the fraction of ground covered by green vegetation, which quantifies the spatial extent of the vegetation. The FCover estimates are obtained through a near real-time algorithm that uses top-of-canopy reflectance observations from the SPOT/VEGETATION satellite data (Verger et al., 2014). The product is available globally at 1 km spatial resolution on day 10, 20 and the last day of each month for the period 1999–2018.

The ESA-CCI Land Cover dataset provides globally consistent maps at 300 m spatial resolution on an annual basis from 1992 to 2015. The land cover typology was based on the Land Cover Classification System (LCCS) developed by the United Nations (UN) Food and Agriculture Organization. A total of 22 land cover level 1 classes and 14 level 2 sub-classes (defined using a set of classifiers) constitute the dataset (Appendix 6.1, Poulter et al., 2015). In this study we used the global map for the year 2010.

2.2 - Methods:

Since the LST is a clear-sky product (it can only be measured when the sky is not cloudy), a threshold was chosen to be representative of clear-sky days in order to allow the comparison between LST and SKT. Only the data meeting the following thresholds was retained in the subsequent analysis:

- The reanalysis's TCC < 0.3;
- The percentage of valid LST original data in each 0.25º × 0.25º grid cell > 0.7.

The two previous thresholds were chosen based on the average percentage of valid data available for the comparison while keeping, at the same time, most of the grid cell cloud-free (Appendix 6.2).

The clear-sky threshold of each ERA reanalysis was also applied to the corresponding HTESSEL offline simulation (forced by that reanalysis) so that all datasets could be compared. The offline simulation forced by ERA-Interim (that has an hourly output) was analysed using a three-hourly time step in order to match with ERA-Interim's three-hourly output.

The analysis was performed between 2004 and 2015, the total LST reprocessed data period available at the time, over the Iberian Peninsula (35ºN–45ºN, 10ºW–5ºE). We only examined the summer months (June-July-August, JJA) due to the overall low percentage of valid data during the rest of the year (e.g. in 2010, before May and after September, the valid data was below 50%). Furthermore, previous studies show that the mean error between reanalyses and observations is higher in the summer (e.g. Trigo et al., 2015; Zhou et al., 2017).

An analysis was carried out to separate the domain into different clusters. The K-Means Clustering Algorithm was the method selected due to its simplicity and fast convergence to a solution. The K-Means is a partitional method in which each datapoint belongs to one cluster only (each cluster being comprised of datapoints with similar characteristics, defined by the input data). The clusters are each represented by a centroid with a real location in the domain. The L2 distance or Euclidean distance (the sum of squared distances in each axis) is computed by the model between each datapoint and all the

centroids. The datapoints are assigned to the cluster with its centroid closest to the datapoints and the mean of each cluster becomes the new centroid. The L2 distance is computed again and the new clusters obtained until the algorithm converges to a solution that doesn't change with further iterations. Also, if the input data is a combination of variables with different units, the data needs to be normalized in order for the L2 distance to be computed correctly.

The K-Means Algorithm calculates the value of inertia (defined as the sum of the L2 distance of the datapoints to their closest cluster centroid), which decreases as the number of clusters increase. The total number of clusters is defined by the user. To ensure that the correct number of clusters was selected, the Elbow Method was applied: the optimal number of clusters (that should be designated into the K-Means algorithm) is that which the addition of an extra cluster would result in a negligible change in inertia (a decrease of less than 10% of its value). Nevertheless, this method presents a few drawbacks, like its tendency to converge to local minima and its dependence on the number of clusters selected.

A set of metrics was chosen to assess the quality of the four different products analysed in this work (see Table 2.3):

1. **Mean Error (or Bias)**. It represents the difference in the mean value of a variable between the model and the reference product.
2. **Standard Deviation of the Error** (SD). It gives information about the random variability in a given product.
3. **Temporal Correlation**. Measured with the Pearson's correlation coefficient, *r*. It characterizes the temporal coincidence of variability.
4. **Root Mean Squared Error** (RMSE). It is an aggregation of the information given by the Mean Error and Standard Deviation of the Error.

Each metric was calculated for the daily maximum (Tmax) and the daily minimum temperature (Tmin). Tmax and Tmin were calculated for the whole domain and for the whole period in the following range of hours:

- Tmax: 11h – 15h;
- Tmin: 3h – 7h.

These ranges of hours were chosen based on the assumption that the sky is always clear in the data used in this study, after computing the clear-sky threshold, hence the maximum temperature will occur shortly after the peak of incoming solar radiation, while the minimum temperature will precede the sunrise.

The four metrics were applied for the whole domain and the results were then grouped into the clusters identified by the K-Means algorithm.

To conclude the first part of this work, we assessed the relation between the error in the simulation of SKT and the error in the representation of vegetation in the HTESSEL model. The HTESSEL fraction of vegetation cover (HTESSEL-FCover) was compared to a mean value of the observations (FCover) from CGLS over Iberia. HTESSEL-FCover has a constant value in each grid point, given by Equation (2.1), while CGLS-FCover was computed as the 1999-2018 mean value in each pixel.

Table 2.3. Performance metrics used in the evaluation of the products. P stands for the Product being assessed and R is the reference dataset (in this case, the satellite-LST).

Mean Error	$ME = \frac{\sum_{i=1}^{n} P_i - R_i}{n}$
Standard Deviation of the Error	$SD = \sqrt{\frac{\sum_{i=1}^{n}[(P_i - R_i) - (\bar{P} - \bar{R})]^2}{n-1}}$
Temporal Correlation	$r = \frac{n\sum_{i=1}^{n} P_i R_i - \sum_{i=1}^{n} P_i \sum_{i=1}^{n} R_i}{\sqrt{\left[n\sum_{i=1}^{n} P_i^2 - \left(\sum_{i=1}^{n} P_i\right)^2\right]\left[n\sum_{i=1}^{n} R_i^2 - \left(\sum_{i=1}^{n} R_i\right)^2\right]}}$
Root Mean Squared Error	$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - R_i)^2}{n}}$

In the second part of the thesis, we tried to improve the simulation of SKT by the HTESSEL model. The domain comprised the four grid points surrounding Évora (38.25ºN–38.75ºN, 8.25ºW–7.75ºW) with the same 0.25º × 0.25º resolution as the original Iberia domain. This domain was initially selected as the surroundings of a station used to validate the LSA-SAF satellite LST retrievals, but it was not possible to process the point observation for this work. Despite this, the area chosen is representative of the main features and errors explored. The results were computed for the summer of 2010 (whose scores were comparable to the summers of 2004-2015) to reduce the computational cost and data handling.

We started out with three simplified simulations to assess the influence of the original modelled vegetation in the SKT output: the bare soil in 'bare', where CVL and CVH were assigned to zero (in other words, the domain is assumed to be a desert); the low vegetation in 'lveg', where CVL original value was kept and CVH was changed to zero; and the high vegetation in 'hveg', where CVL was altered to zero while CVH kept its original value. The corresponding FCover was obtained using Equation (2.2).

These simple tests were followed by two other experimental simulations, this time using the CGLS-FCover: 'nlveg', where CVH was changed to zero and CVL was given the value of the CGLS-FCover, and 'nhveg', where the CGLS-FCover was attributed to CVH while CVL was altered to zero. c_{veg} was assumed to be 1 (the vegetation covers its fraction of the domain in its entirety) so that FCover is equal to CVL in 'nlveg' and CVH in 'nhveg'. The CGLS-FCover considered was the mean of the year 2010.

The previous experiments are summarized in Table 2.4.

After these simplified simulations, we decided to perform a more elaborate simulation by introducing a revised vegetation in the model. The ESA-CCI Land Cover dataset was converted to Plant Functional Types (PFTs) using the 'cross-walking' table suggested by Poulter et al. (2015) (Appendix 6.1). PFTs represent groups of plant species with similar characteristics (structural, phenological and physiological, and their climate zones) to reduce the complexity of the vegetation representation in models (Poulter et al., 2015). However, the HTESSEL lookup table (Table 2.1) is not identical to the PFTs. Instead of performing *ad hoc* matching between the PFTs and HTESSEL vegetation types, we kept the model TVL and TVH, assuming their accuracy, and changed the model CVL and CVH to the values obtained with the ESA-CCI dataset.

Table 2.4. Setup of the experiments employed in this study.

Experiment\Parameter	CVL	CVH	c_{veg}	FCover
bare	0	0	Given by Table 2.1	Obtained with Equation (2.2)
lveg	Original value	0		
hveg	0	Original value		
nlveg	CGLS-FCover	0	1	CGLS-FCover
nhveg	0	CGLS-FCover		
revised	ESA-CCI		Table 2.1	Equation (2.2)

Additionally, we altered the soil discretization (from four to nine layers) and assessed its impact in the control and in the revised simulations.

In the last part of the thesis, we performed a sensitivity study on the c_{veg} parameter to explore its influence in the SKT output and subsequent error. This analysis was performed with the original and the revised modelled vegetation. We perturbed c_{veg} associated with the TVL and TVH considered by the HTESSEL model, within the range 0.1–1, forming 100 perturbations composed of c_{veg} pairs. The pairs were determined with a quasi-random sampling method, the Sobol sequences (Sobol, 1967), which are designed to efficiently generate samples of the multiple parameters that cover the entire parameter space while avoiding the introduction of correlations between the perturbations of the different parameters. Unlike in random sampling, the sample values in the Sobol approach are selected based on the previously generated values to prevent the occurrence of clusters or empty spaces in the domain.

3 - Results and Discussion

3.1 - Evaluation

The percentage of valid data in the summer in the period of 2004-2015 (after applying the clear-sky threshold) is represented in Figure 3.1. Most of the Iberian Peninsula has over 50% of available data in the 12-year period.

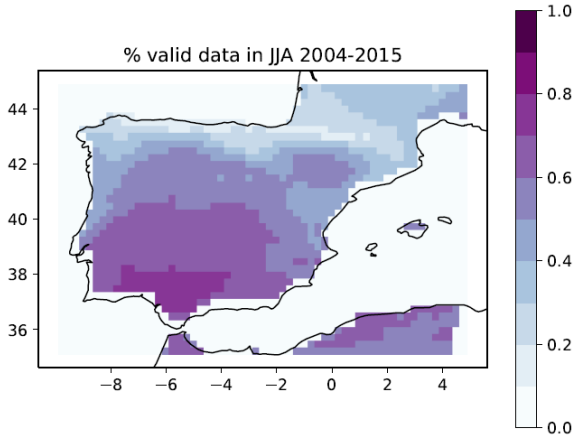


Figure 3.1. Percentage of valid data in the summer between 2004 and 2015.

The K-Means clustering algorithm grouped the datapoints into six different clusters (Figure 3.2). The input data was the maximum and minimum LST of the mean diurnal cycle in each pixel in the summer months for the period 2004-2015. LST was selected as the input temperature since it is considered the reference temperature in this work. The six clusters represented regions with different LST diurnal cycles (Figure 3.2), although some of the regions show similar diurnal cycles for both ERA reanalyses.

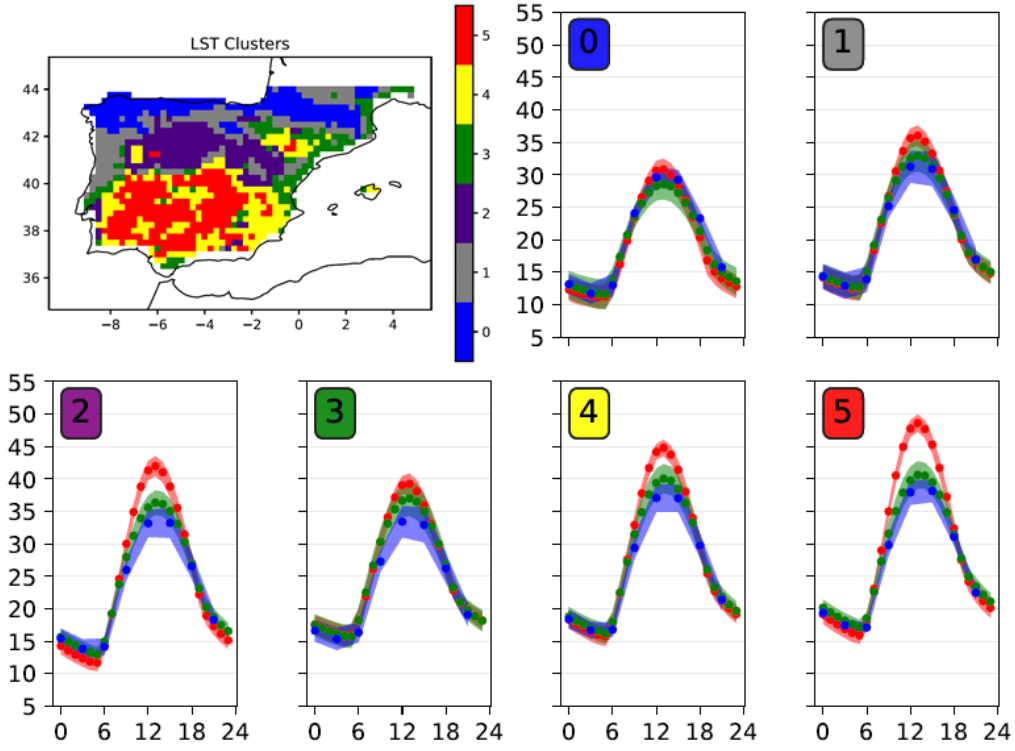


Figure 3.2. Clusters determined by the K-Means Algorithm with LST as input (top left) and the average (dots) and SD (shaded) 2004-2015 JJA diurnal cycle (℃) in each cluster for the satellite-LST (red), ERA5 (green) and ERA-Interim (blue).

We tested the K-Means Algorithm with different input data, namely the CGLS-FCover first and, afterwards, both LST and FCover data. For the LST+FCover clusters, the LST data was normalized by simply dividing the data by the maximum temperature in the dataset (since there were no sub-zero temperatures in the dataset). In general, the results were similar but the FCover 'inland' clusters (clusters 3 and 4 in Appendix 6.3) differed from the LST 'inland' clusters (clusters 2, 3 and 4 in Figure 3.2): there was no separation between the areas to the north and to the south of the Central System in the FCover clusters. While the vegetation cover is similar in both areas (Appendix 6.5), the area to the north of the Central System, the Iberian Plateau, is located at a higher altitude than the area to the south (Appendix 6.6). The LST data includes the land cover signal but also accounts for the orography effect, as seen by the LST+FCover clusters in Appendix 6.4 which are very similar to the clusters obtained using only the LST data (Figure 3.2).

Figures 3.3 to 3.6 show the within-cluster distributions of the Tmax Bias, SD and RMSE, and Tmin RMSE by means of boxplots. The Tmax Bias is consistently negative in the 'inland' clusters of Figure 3.2, with ERA-Interim showing slightly larger bias than ERA5 (Figure 3.3). Both surface experiments have similar biases to ERA5, which shows that the updates in the HTESSEL model allow for a better simulation of SKT, even if its atmospheric forcing is of reduced quality (like in the case of the simulation forced by ERA-Interim, SEI). The Tmin Bias is overall lower than in Tmax but the results are conspicuous (Appendix 6.7). The simulation forced by ERA5 (SE5) is systematically colder than ERA5, while ERA-Interim and SEI are closer and with slightly smaller errors than ERA5.

The Tmax SD is around 2–3 K in general, with higher values in ERA-Interim and SEI (Figure 3.4). This might be explained by the better quality of the meteorology dynamics in ERA5 (e.g. high spatial resolution) also present in SE5. The Tmin SD is in general smaller than Tmax (with values between 1–2 K). Although ERA5 shows better results than ERA-Interim overall (and SE5 in relation to SEI), the surface experiments show a lower performance (Appendix 6.8). This inconsistency in performance between the offline simulations and the original reanalysis for Tmax and Tmin is likely related to the representation of stable conditions for Tmin which does not reproduce exactly the coupled system, associated with the stability functions of the exchange coefficients and differences in timestep.

On the other hand, the temporal correlation showed very similar values among the four products, for both Tmax and Tmin (Appendices 6.9 and 6.10). Still, it is interesting to notice that, for this metric, ERA-Interim and SEI have consistently better correlations than ERA5 and SE5 in Tmax. The coarser temporal and spatial resolution of ERA-Interim might explain these results.

The 'inland' clusters show a higher RMSE in Tmax (median above 4 K) (Figure 3.5). Overall, ERA5 and both HTESSEL simulations have lower errors than ERA-Interim. The RMSE in Tmin (Figure 3.6) is much lower than Tmax (median value within 1 K and 2 K), following the other metrics' results.

In Figures 3.7Figure 3.7 and 3.8 the Tmax and Tmin RMSE are represented, respectively, for the whole domain. The zones with higher Tmax RMSE are similar among the four products, comprising the SW Iberian region and the northern Iberian Plateau, where RMSE $>$ 10 K in a vast number of grid points. The Tmin RMSE is overall higher in mountainous regions although the zones with higher RMSE differ in each product.

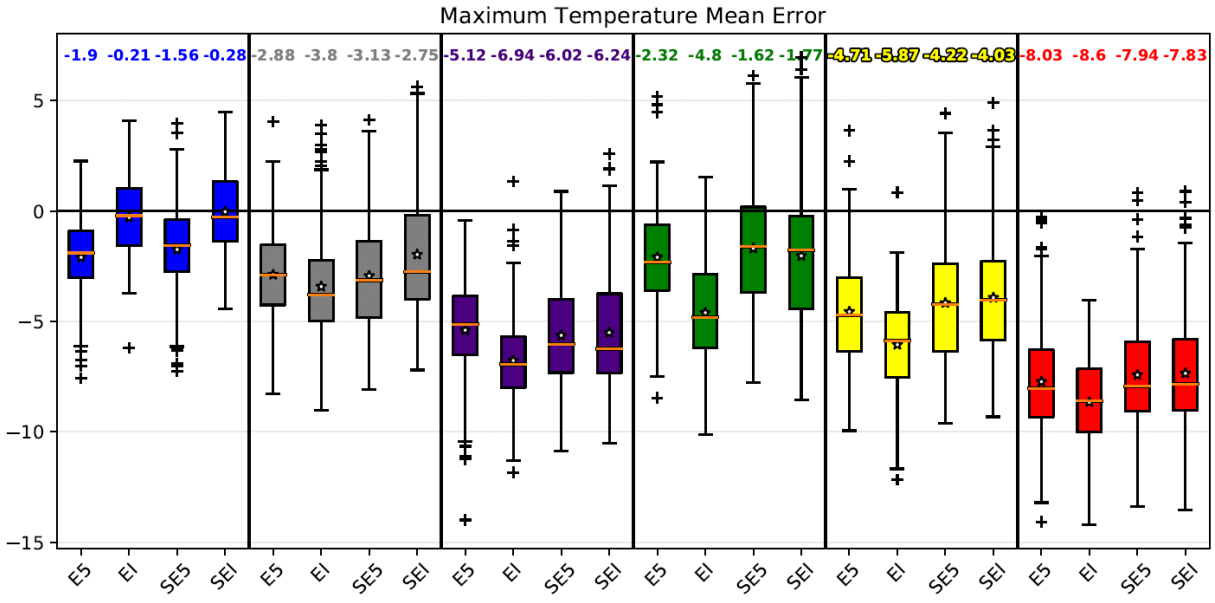


Figure 3.3. Tmax Bias (K) in: ERA5 (E5), ERA-Interim (EI) and the simulations forced with E5 (SE5) and EI (SEI). The colours represent the clusters in Figure 3.2. The number above each boxplot is the median of that boxplot.

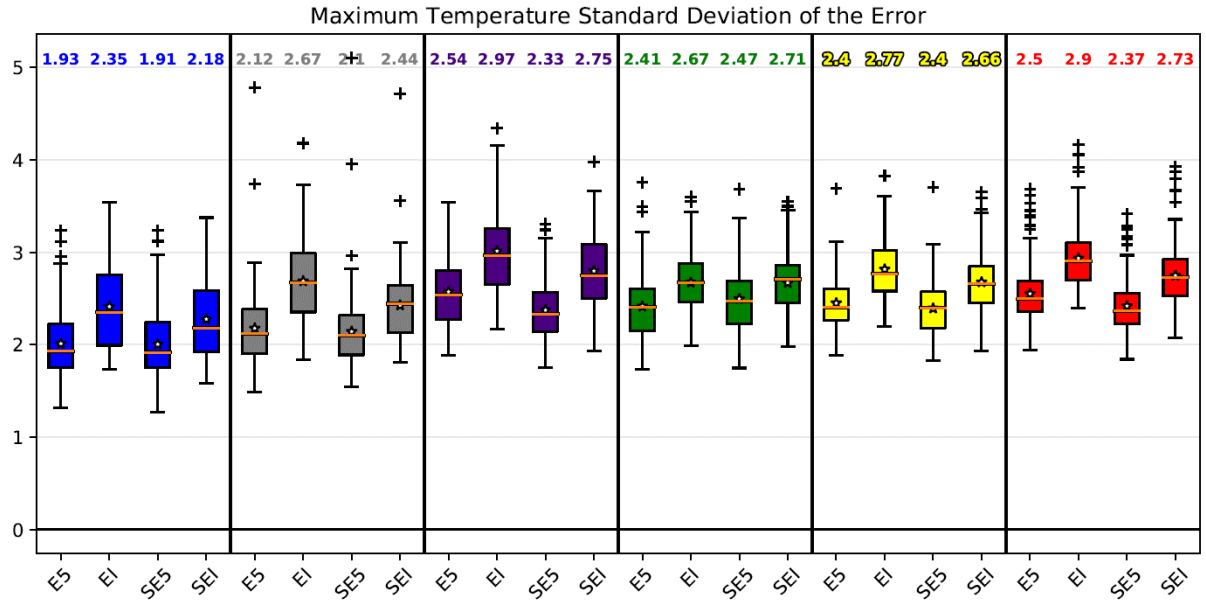


Figure 3.4. Tmax SD (K) in: ERA5 (E5), ERA-Interim (EI) and the simulations forced with E5 (SE5) and EI (SEI). The colours represent the clusters in Figure 3.2. The number above each boxplot is the median of that boxplot.

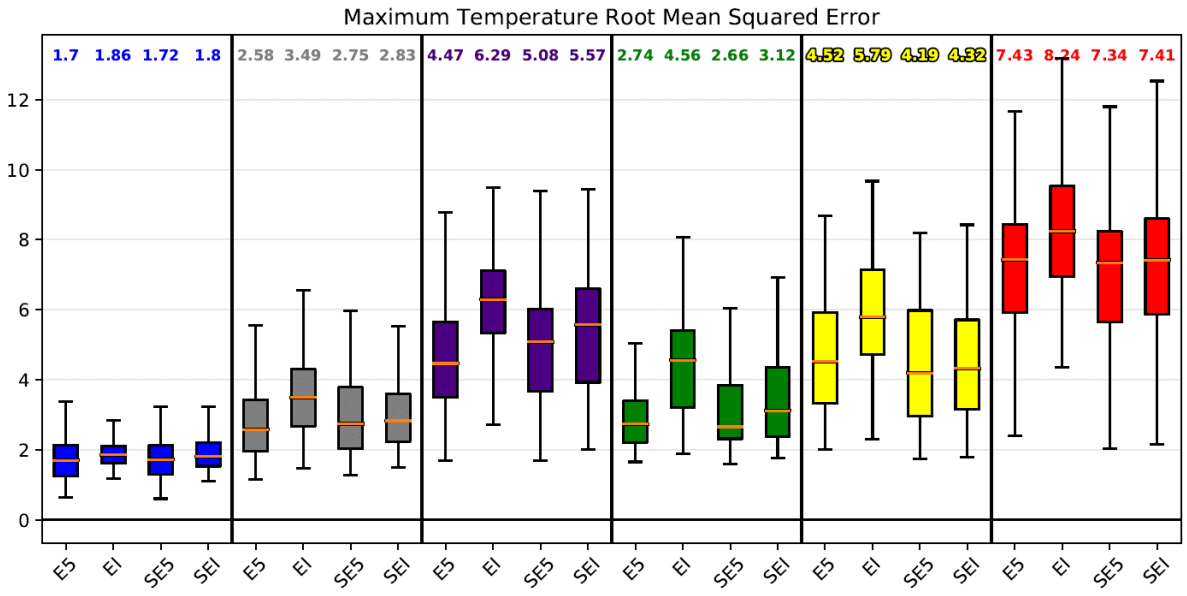


Figure 3.5. Tmax RMSE (K) in: ERA5 (E5), ERA-Interim (EI) and the simulations forced with E5 (SE5) and EI (SEI). The colours represent the clusters in Figure 3.2. The number above each boxplot is the median of that boxplot.

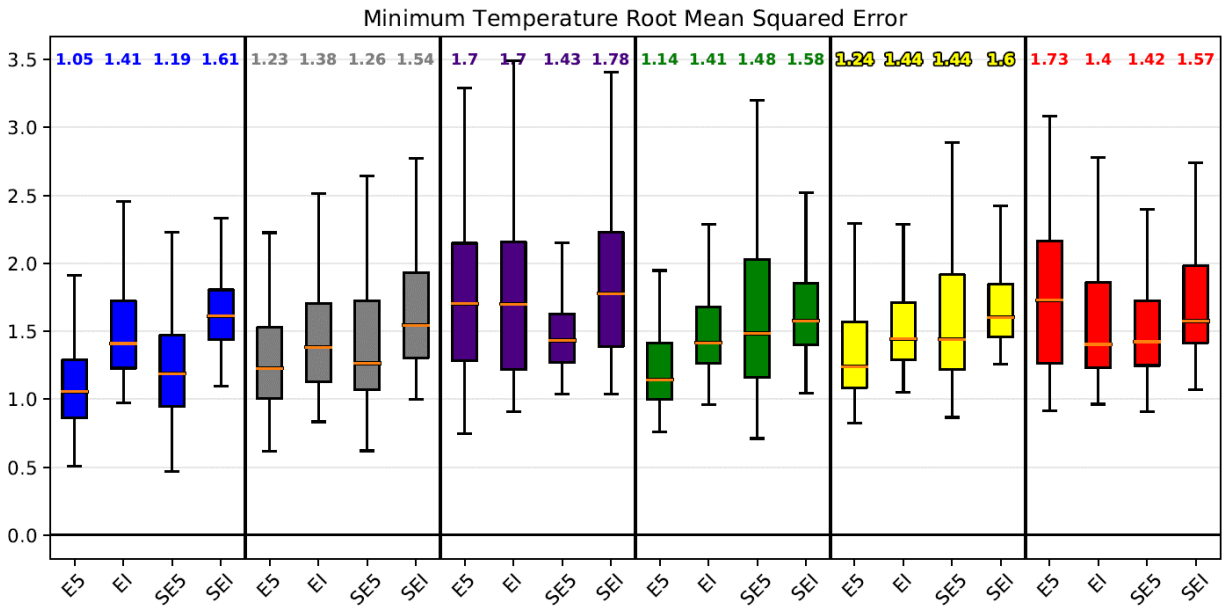


Figure 3.6. Tmin RMSE (K) in: ERA5 (E5), ERA-Interim (EI) and the simulations forced with E5 (SE5) and EI (SEI). The colours represent the clusters in Figure 3.2. The number above each boxplot is the median of that boxplot.

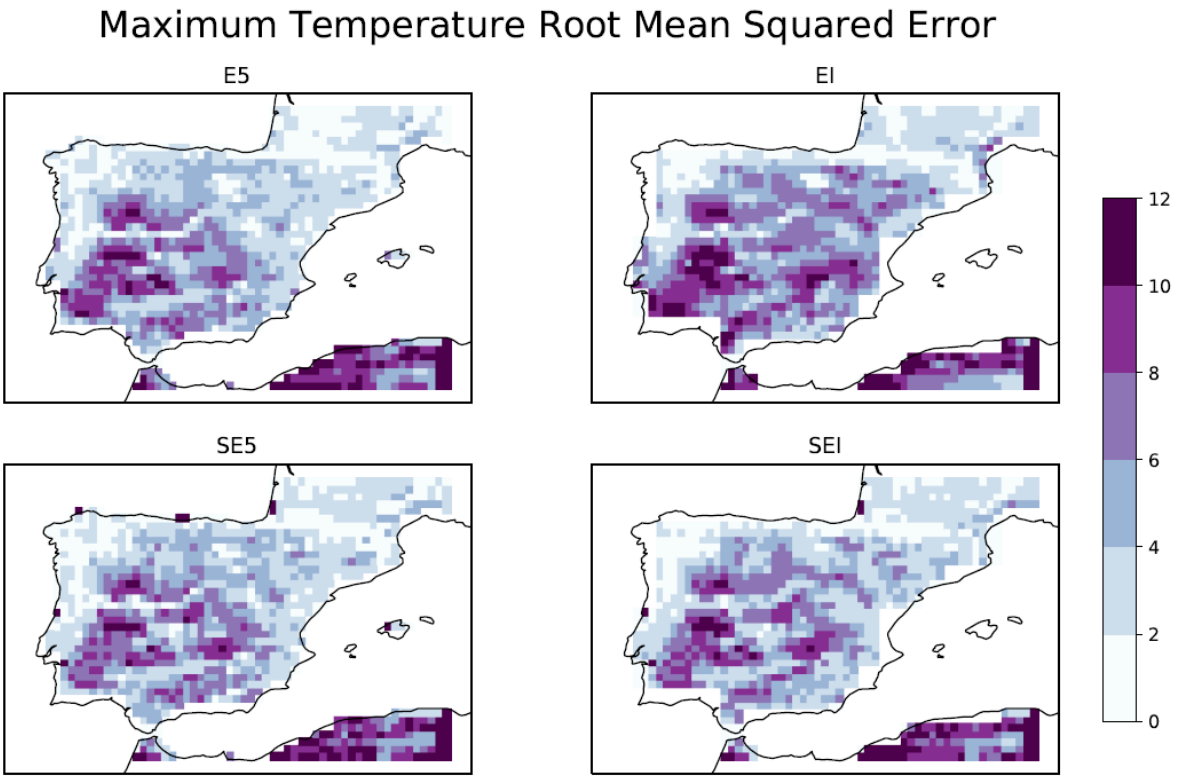


Figure 3.7. Tmax RMSE (K) in ERA5 (E5), ERA-Interim (EI) and the simulations forced with E5 (SE5) and EI (SEI).

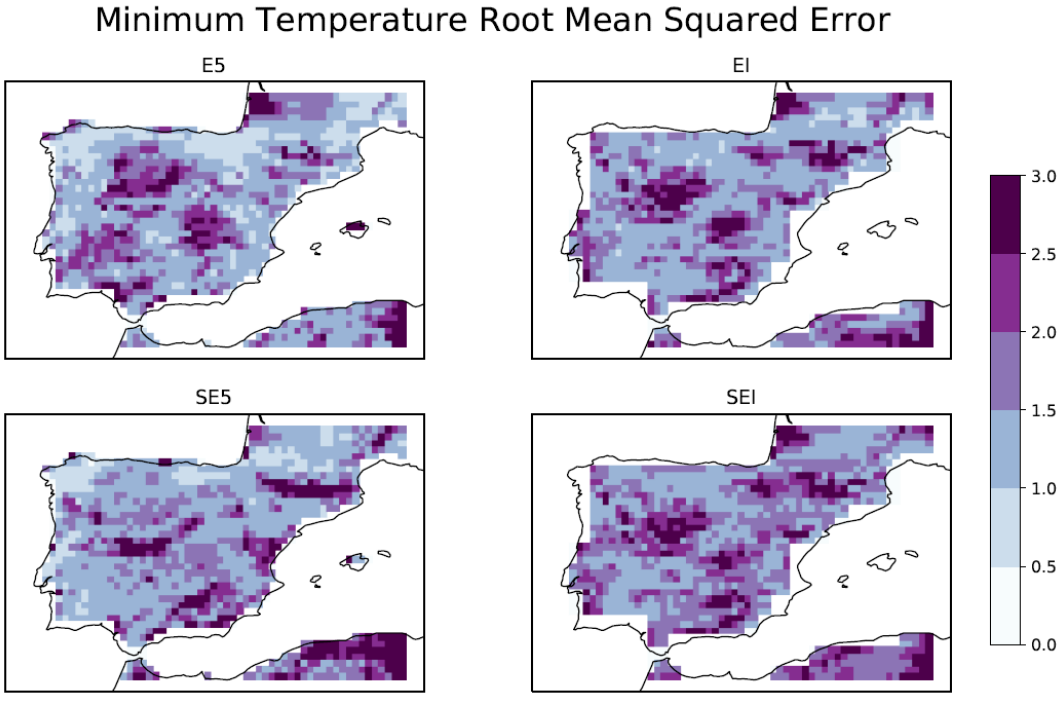


Figure 3.8. Tmin RMSE (K) in ERA5 (E5), ERA-Interim (EI) and the simulations forced with E5 (SE5) and EI (SEI).

In general, the evaluation results are similar to the ones presented in previous works (e.g. Trigo et al., 2015; Orth et al., 2017; Zhou et al., 2017), with strong daytime underestimation and a weak nighttime overestimation of SKT in the summer, especially in semi-arid regions.

This analysis also allows to show that the offline (uncoupled) simulations have similar errors to the reanalyses and can therefore be used to apply changes in model parameters and to assess their impact in the SKT simulation.

The clear-sky threshold chosen has an impact on the results, although if a more restrictive threshold had been chosen, that would have resulted in a lesser percentage of data meeting the threshold while the associated cloud cover values would have remained similar (low values overall; Appendix 6.2).

Finally, Figure 3.9 presents a positive correlation (0.45) between the difference of the HTESSEL FCover with respect to the CGLS and the Tmax Mean Error (Appendix 6.5 shows the CGLS- and the HTESSEL-FCover in Iberia).

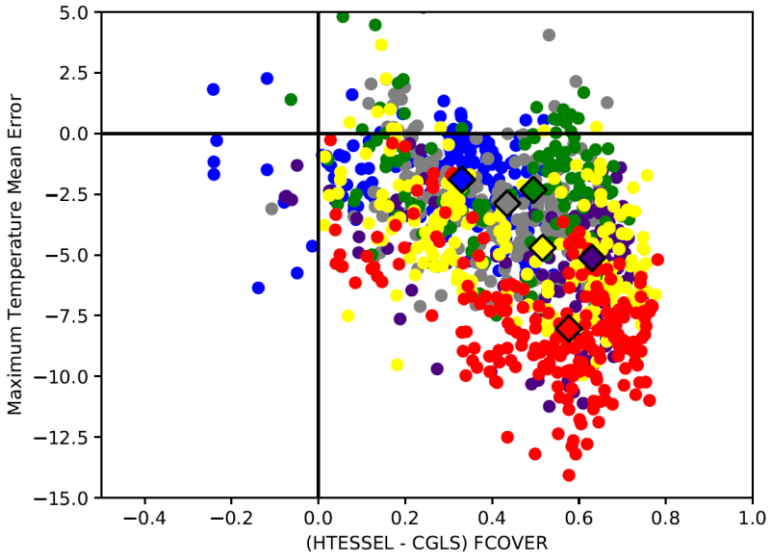


Figure 3.9. Tmax Mean Error (K) in ERA5 (y-axis) and the CGLS-FCover error in the HTESSEL model (x-axis). The colours represent the clusters in Figure 3.1. The diamond symbol represents the median value of each cluster.

A possible next step involved attempts to enhance the representation of vegetation in the HTESSEL model.

3.2 - Simulations with revised parameters

The first parameter to be reviewed was the model representation of vegetation cover. The model's original TVL, TVH, CVL and CVH in the Évora-centred domain (Figure 3.10) is presented in Table 3.1, and we determined the associated HTESSEL-FCover using Equation (2.2.

In all points, the dominant high vegetation type is 'Interrupted Forest' while the dominant low vegetation type is 'Evergreen Shrubs' (except for the SW point), with most of the grid point covered by high vegetation ($CVH > 0.80$ except for the NW point).

Table 3.1. Original model vegetation parameters in the four-point domain centred in Évora (Figure 3.10).

Parameter \ Point	NW	NE	SW	SE
TVL	16	16	0*	16
TVH	19	19	19	19
CVL	0.40	0.07	0.01	0.17
CVH	0.51	0.93	0.99	0.81
FCover (HTESSEL)	0.66	0.87	0.89	0.81

* As default, 0 is assigned to index 8 – Desert (that is, the low vegetation in the SW point is negligible)

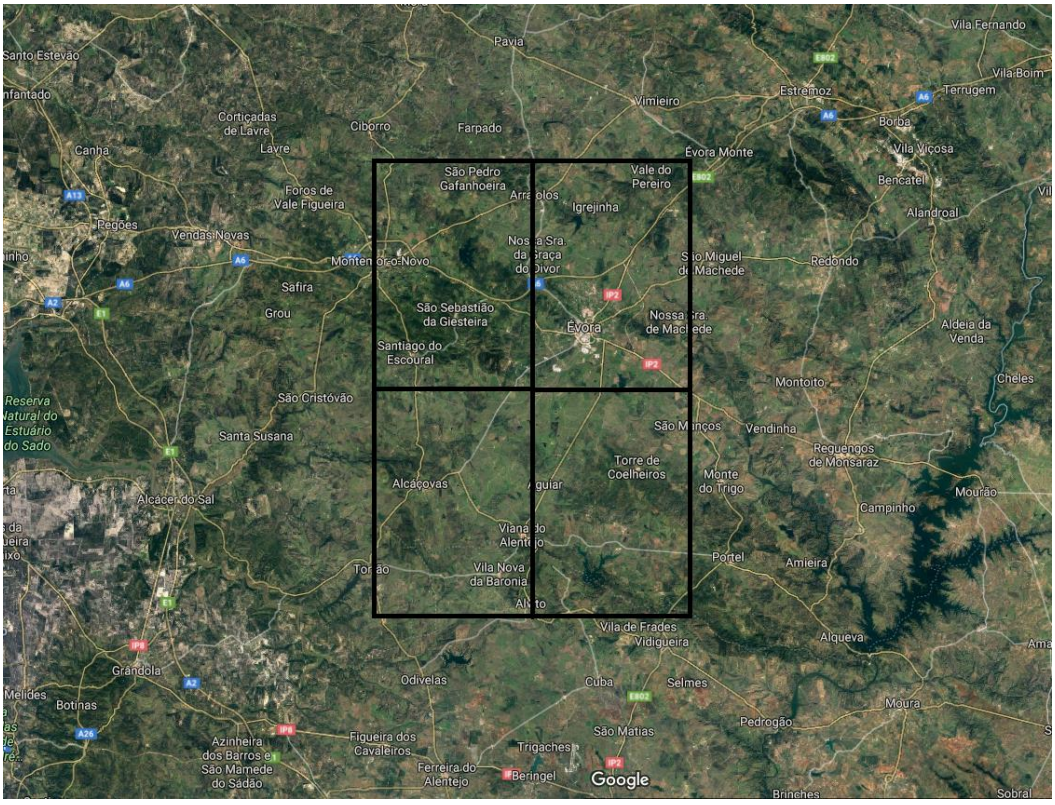


Figure 3.10. Satellite view of the approximate Évora-centred domain.

The SKT diurnal cycle in the summer of 2010 of the 'control' simulation presents strong cold biases during the day and weak warm biases at night (Figure 3.11). The NE and SW points have the warmest satellite-LST diurnal cycles while the NW and SE points have the warmest 'control' SKT diurnal cycles. As a result, the NW point (where the HTESSEL-FCover is much lower (0.66) in relation to the remaining points) shows the best approximation of the model simulation to the satellite observations (with a negative bias of 4.5 K at 12 UTC). It is also the only point where the 'bare ground' SKT (the warmest SKT) comes very close to the satellite-LST. The 'high vegetation' SKT is very similar to the 'control' SKT in all the points while the 'low vegetation' SKT has the lowest values, which is attributed to CVH and CVL having values close to one and zero, respectively. The 'control' SKT diurnal cycle presents a slight phase difference in relation to the satellite-LST during the day, taking longer to warm in the morning and to cool down in the afternoon.

Additionally, in Figure 3.11 the temperature of the first soil layer and the T2m are represented. T2m is very similar in all the points, with much lower values than the SKT during the day and slightly higher values at night. The first soil layer temperature is highest in the NW point (where the 'control' SKT is also highest) and follows a similar diurnal cycle to the T2m's, but with higher temperatures overall. The '9 soil layers' experiment did not significantly change the 'control' SKT (dashed green line), leading to a bias decrease of ~ 0.5 K, but it shows a reasonable variation in the first soil layer temperature (dashed brown line).

The percentage of valid data in the domain in the summer of 2010 varied between 0.70 and 0.73, which is similar to the 2004-2015 JJA mean in Figure 3.1.

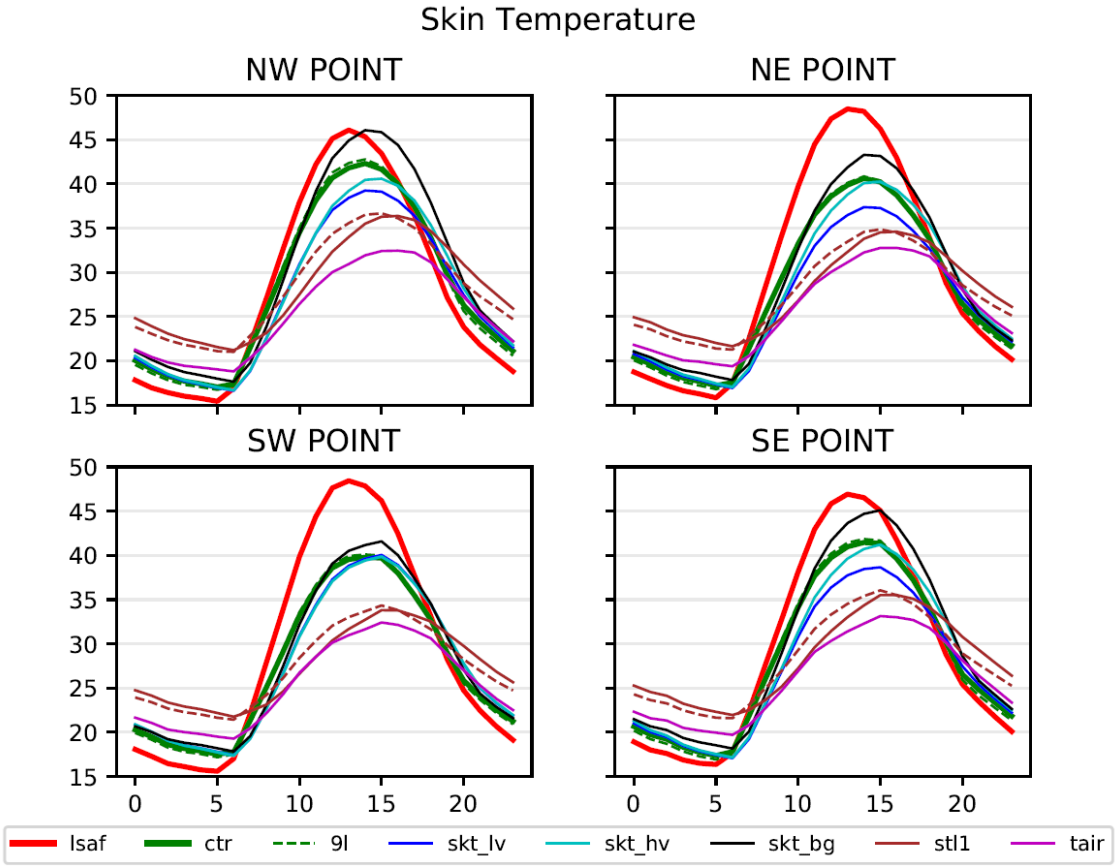


Figure 3.11. 2010 JJA Mean Diurnal Cycle of SKT (°C) in the Évora-centred domain: satellite-LST (red), control (green) and 9 soil layers ('9sl'; dashed lines) simulations. 'skt_lv', 'skt_hv' and 'skt_bg' represent, respectively, the SKT diurnal cycle of low vegetation-, high vegetation-, and bare ground-only tiles. T2m (tair; pink) and the temperature of the first soil layer (stl1; brown) are also plotted.

The results of the simplified simulations are presented in Appendix 6.11. The two experiments with CGLS-FCover ('nlveg' and 'nhveg') present a diurnal cycle much closer to the satellite observations, in particular the 'nlveg' simulation, and seem to hint that the actual value of the FCover might be much lower than the given modelled value. The simplified simulations with the original model parameters ('bare', 'lveg' and 'hveg') provide a similar conclusion as well: the 'bare' and 'lveg' (the 'hveg' and 'control') simulations are very similar to each other due to CVL (CVH) being close to zero (one) in all the points except for the NW one, with all experiments remaining distant to the satellite-LST. The 'nlveg' and 'nhveg' experiments also suggest that the low vegetation type might be the main type in the domain (that is, $CVL > CVH$). In connection with the previous statement, changing the TVL in the SW point from 'Desert' to 'Evergreen Shrubs' further approximates the 'nlveg' SKT simulation to the satellite-LST (Appendix 6.12).

The revised vegetation parameters obtained with the ESA-CCI Land Cover dataset are presented in Table 3.2. The original TVL and TVH were kept, except for the SW point where TVL was changed to 16 (Evergreen Shrubs) since CVL is no longer negligible in that point (and this way, the SW point becomes identical to the other three grid points).

Table 3.2. Revised model vegetation parameters in the four-point domain centred in Évora. Values between brackets denote the original model vegetation parameters as in Table 3.1 for comparison purposes.

Parameter \ Point	NW	NE	SW	SE
CVL	0.87 (0.40)	0.89 (0.07)	0.89 (0.01)	0.85 (0.17)
CVH	0.13 (0.51)	0.08 (0.93)	0.10 (0.99)	0.12 (0.81)
FCover (HTESSEL)	0.55 (0.66)	0.52 (0.87)	0.54 (0.89)	0.53 (0.81)
FCover (CGLS)	0.42	0.41	0.47	0.40

The revised vegetation shows a stark contrast to the original model vegetation. Overall, the HTESSEL-FCover is much lower (~ 0.53) and the grid points are covered mostly by low vegetation ($CVL > 0.85$) when using the ESA-CCI dataset. When comparing HTESSEL's revised FCover to CGLS's, even though CGLS's is lower in all the grid points (between 0.07 and 0.13 than the revised HTESSEL-FCover), they are both closer to each other than to the 'control' HTESSEL-FCover. These results follow the preliminary findings with the simplified experiments.

It is important to mention that the ESA-CCI's TVL is not a direct match to the HTESSEL's TVL, as seen in Appendix 6.13. In the Évora-centred domain, HTESSEL's TVL is "Evergreen Shrubs" while in ESA-CCI's PFTs is "Managed Grass". It is not clear which TVL in HTESSEL is the equivalent (or the closest) to "Managed Grass" (and for this reason we kept the original TVL when CVL was revised). Experimenting with different TVLs, none produced a diurnal cycle of SKT as close to the satellite-LST as the revised simulation (that kept the original TVL; Appendix 6.14). These results can be partially explained by the c_{veg} used in HTESSEL for "Evergreen Shrubs" (0.5, see Table 2.1) which is lower than the c_{veg} of the remaining TVLs. In other words, we were increasing the FCover when changing the TVL to a different one. The TVH in HTESSEL is "Interrupted Forest" and in ESA-CCI is "Tree Broadleaf Deciduous" (Appendix 6.13). The matching issue does not happen for the TVH because, taking into account Table 2.1, most of the parameters associated with "Interrupted Forest" and "Deciduous Broadleaf Trees" have similar values, so it can be said that the TVH is very similar in HTESSEL and the ESA-CCI dataset.

The enhanced representation of vegetation was shown to have a positive impact on SKT (Figure 3.12). During the day, the SKT becomes very close to the satellite-LST, with a negative bias below 2 K at 12 UTC in all the points, except the NE one. The bias becomes positive in the afternoon (but with an absolute value lower than in the control simulation), as the simulation continues to show a phase difference in relation to the satellite-LST. At night, the impact is almost null.

Parallel to the 'control' simulation, the soil discretization scheme produced a slightly positive effect in the 'revised' simulation, reducing the bias at night by 0.5–1 K and at midday by ~ 1 K (Figure 3.12). These results show that vegetation cover dominates over the soil vertical discretization in terms of SKT simulation.

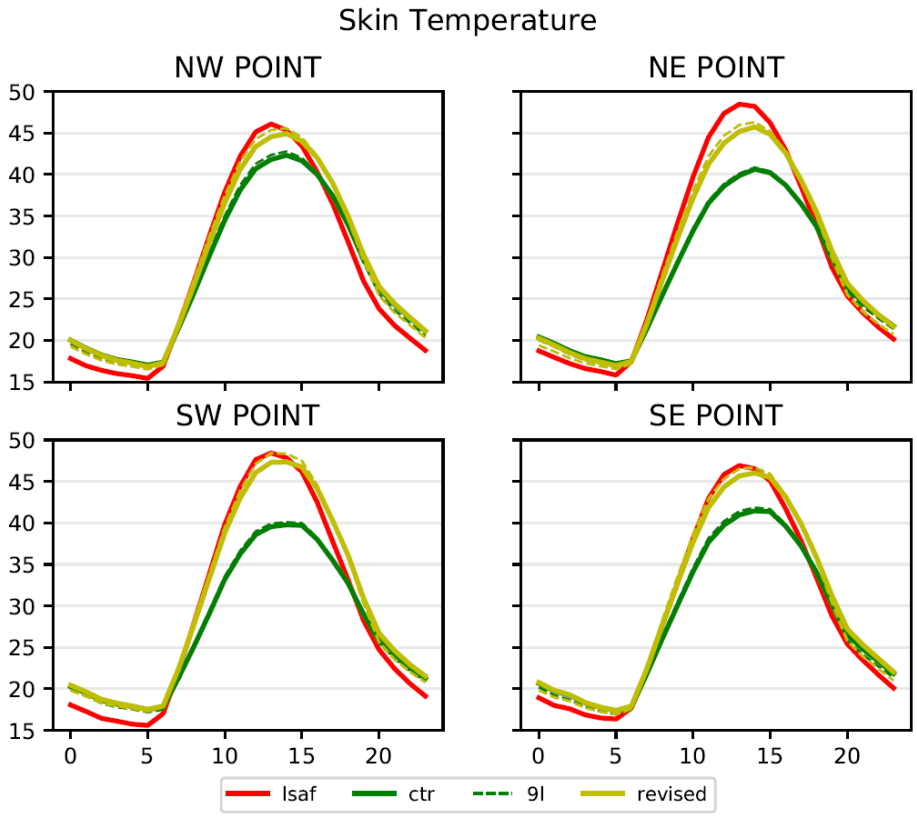


Figure 3.12. 2010 JJA Mean SKT Diurnal Cycle (°C) in the 4-point domain centred in Évora for satellite-LST (red), control (green) and revised (yellow) simulations.

The new representation of vegetation reduces the significant daytime bias that the SKT diurnal cycle exhibited, bringing its value down to the range of the nighttime bias, which in turn remains the same overall. The vegetation plays a major role in determining the SKT during the day but the same cannot be said at night. The nighttime bias is likely related to other model parameterizations that were not assessed in this work. Another issue left to resolve is the existing phase difference in relation to the satellite-LST, especially in the afternoon. Nonetheless, the satellite measurements have their own uncertainty and the error associated with the revised simulation may perhaps lie within that uncertainty's range.

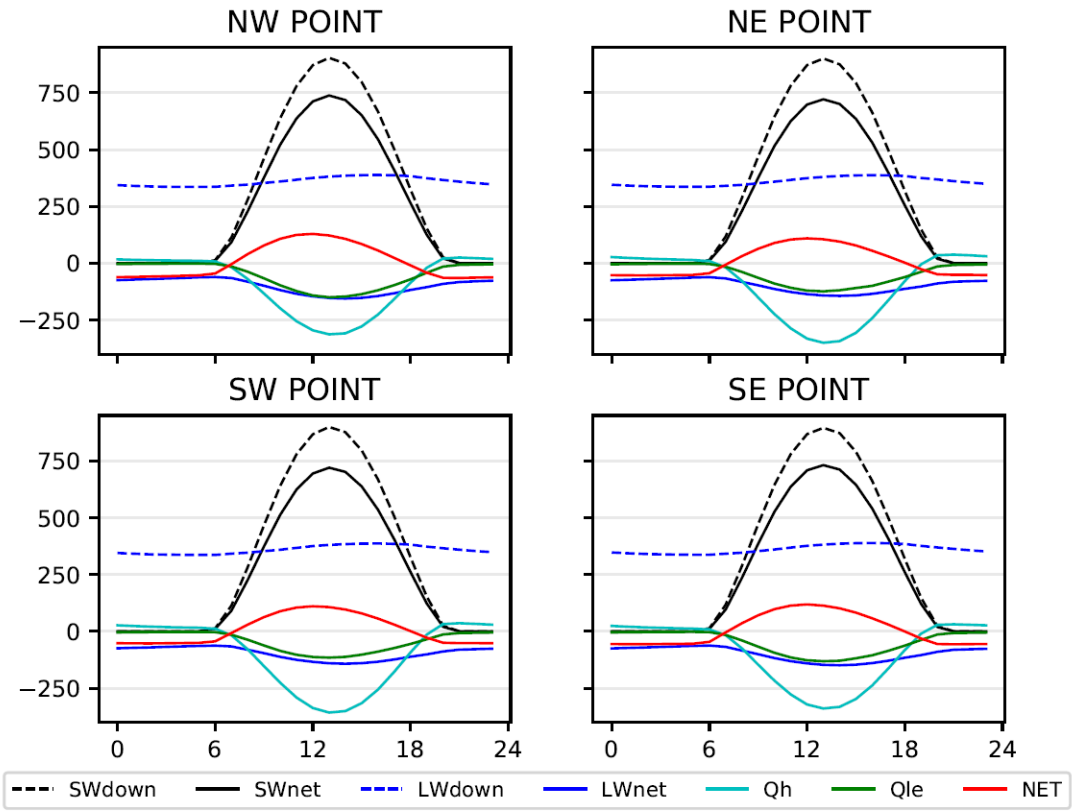


Figure 3.13. Surface energy balance of the control simulation (W/m²): shortwave down (SWdown), shortwave net (SWnet), longwave down (LWdown), longwave net (LWnet) radiation, sensible heat flux (Qh), latent heat flux (Qle) and the net flux (NET=SWnet+LWnet+Qh+Qle).

In Figure 3.13 the components of the surface energy balance of the control simulation are represented. Downward fluxes (to the ground) have positive values and upward fluxes (from the ground) have negative values. Not surprisingly, the surface receives mainly shortwave radiation (SWnet > 0) and emits mainly longwave radiation (LWnet < 0). The sensible (Qh) and latent (Qle) heat fluxes are mostly negative (the heat and moisture transports happen from the ground to the atmosphere), with Qh being slightly positive at night (the heat transport is towards the ground). The net flux (NET) is positive during the day (the surface warms – energy sink of the atmosphere) and negative at night (the surface cools down – energy source to the atmosphere).

A comparison between the energy components of the control and the revised simulations is available in Figure 3.14. At every grid point, the longwave upward radiation (LWup) becomes more negative (the surface emits more radiation) in the revised simulation, because LWup follows the Stefan-Boltzmann's Law ($LWup = \varepsilon\sigma(SKT)^4$) and SKT is higher in the revised simulation. The opposite happens to Qh, as it becomes less negative (less transport of heat from the surface). In the Qh equation (2.5), all the variables remain constant with the revised vegetation, except the SKT and C_H. Since the absolute value of Qh decreases, it means that the change in C_H (in this case, a decrease) impacts Qh more than the increase in SKT. C_H depends on z_{0m} and z_{0h} and these two parameters have lower values in "Evergreen shrubs" than in "Interrupted forest" (Table 2.1). The TVL is dominant over TVH in the revised vegetation which explains the decrease in value of C_H when compared to the original vegetation. The Qle from the surface rises in the revised simulation, except in the SW point where it decreases. In the Qle equation (2.6), in addition to C_H and SKT, r_c also varies (it most likely decreases). The SKT increase results in an exponential increase of q_{sat} due to the Clausius-Clapeyron equation, which in turn leads to

a reduction of Qle. The SW point's distinct behaviour could not be explained within the thesis's duration time as it requires further diagnostics associated with the soil water budget. The net flux absolute value increases as well both during the day and at night (mainly due to the overall reduction of Qh) resulting in the increase in SKT seen in Figure 3.12.

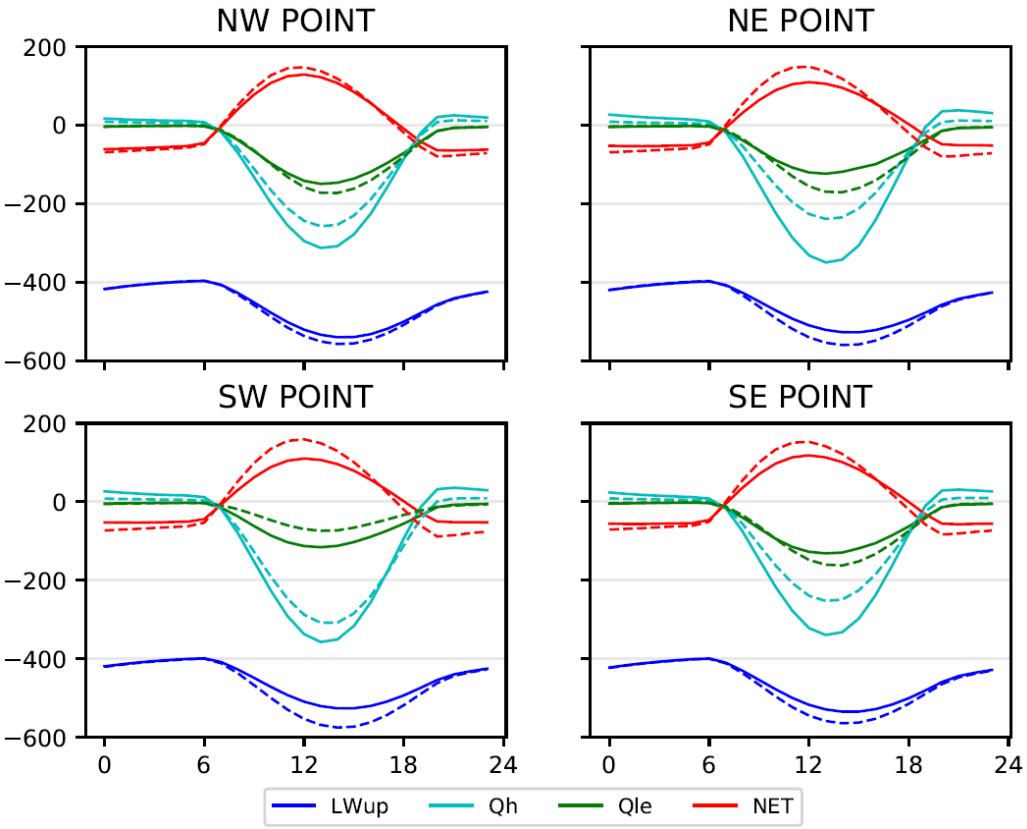


Figure 3.14. Energy Balance at the surface (W/m²): control (solid lines) and revised (dashed lines) simulations.

In summary, in a domain dominated by low vegetation (in this case, shrubs), the surface loses more water (stronger evaporation) and less heat, which warms it.

The uncoupled simulations allowed us to assess the influence of surface parameters in the SKT simulation and the surface energy balance components by varying the value of those parameters (since other factors like the atmospheric variables remain the same in every experiment). Still, it is important to mention that applying these changes in surface parameters in coupled (online) simulations might result in a less positive impact in the simulation of SKT, due to feedback processes associated with the atmospheric coupling (e.g. a smaller r_c might lead to higher wind speed near the surface that, in turn, might cool the earth's surface).

3.3 - Sensitivity study

The sensitivity study applied to the c_{veg} parameter for the original CVL and CVH shows no correlation between the FCover and Tmax Mean Error (Figure 3.16), as the same FCover results in different values of Tmax bias. In Figure 3.15, the perturbations with lower errors have very low values in the high vegetation c_{veg} and the perturbations with higher c_{veg}(H) show a consistent underestimation of Tmax, which indicates once more how the original CVH in the model was too high. The low vegetation c_{veg} can take nearly any value since its influence in the SKT output is very low due to the low CVL in the original model vegetation.

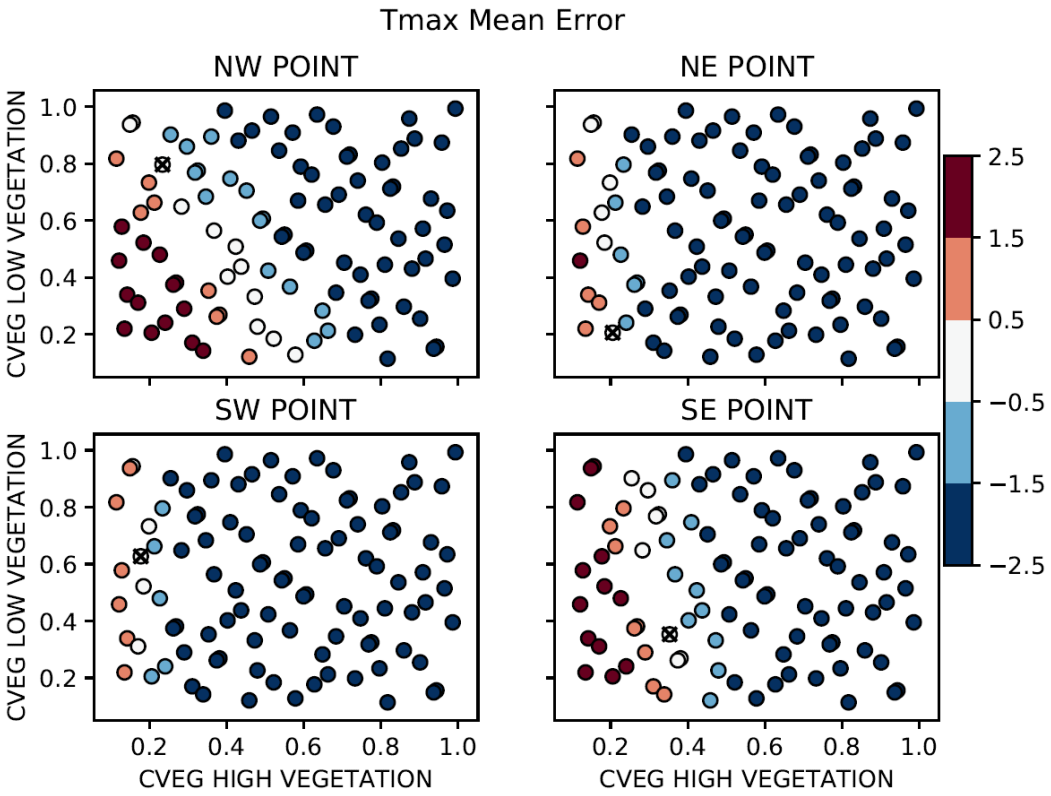


Figure 3.15. c_{veg} perturbation pairs and associated Tmax Mean Error (K) of the original model vegetation. The cross marks the best pair (with Tmax bias closer to zero).

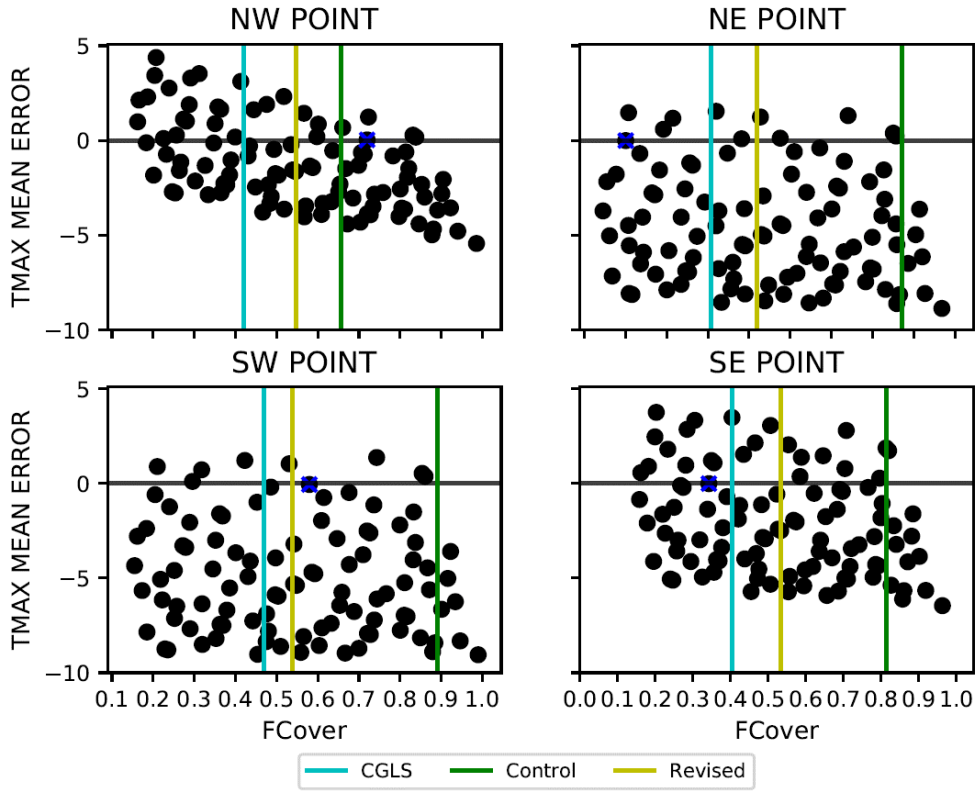


Figure 3.16. Scatterplots of the Tmax Mean Error (y-axis; K) and the FCover (x-axis) of the 100 perturbations (with the original vegetation) in the Évora-centred domain. The blue cross marks the best perturbation (with Tmax Mean Error closer to zero).

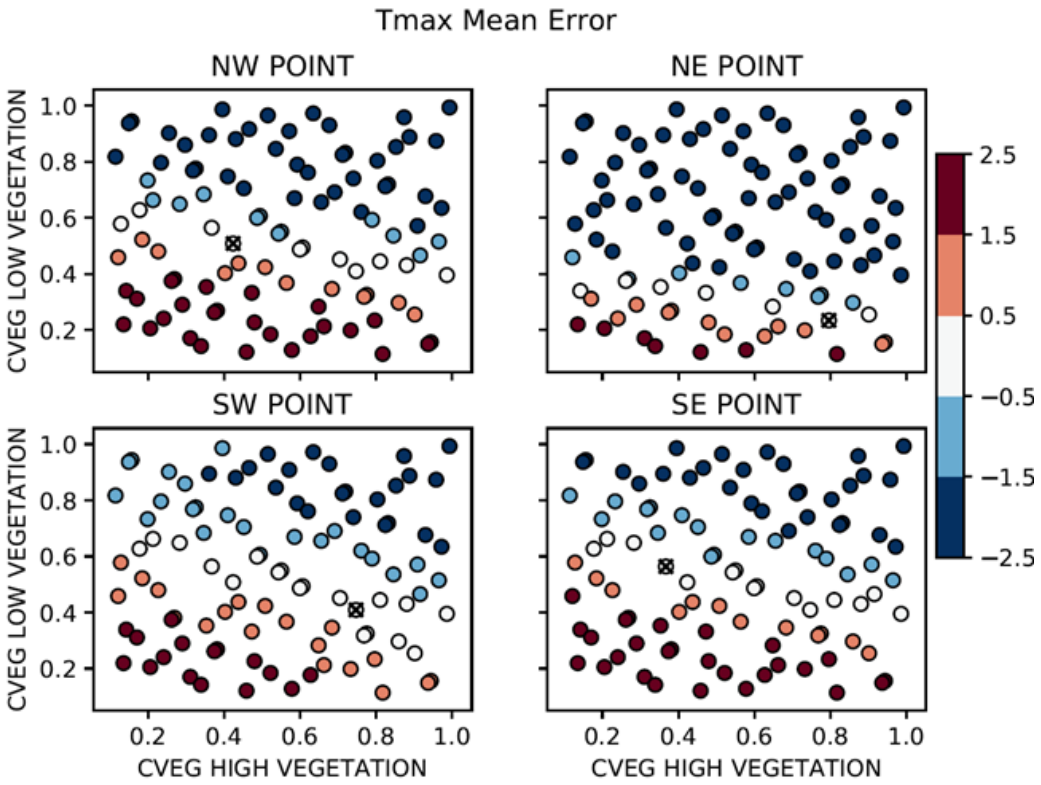


Figure 3.17. c_{veg} perturbations pairs and associated Tmax Mean Error (K) of the revised model vegetation. The cross marks the best pair (with Tmax bias closer to zero).

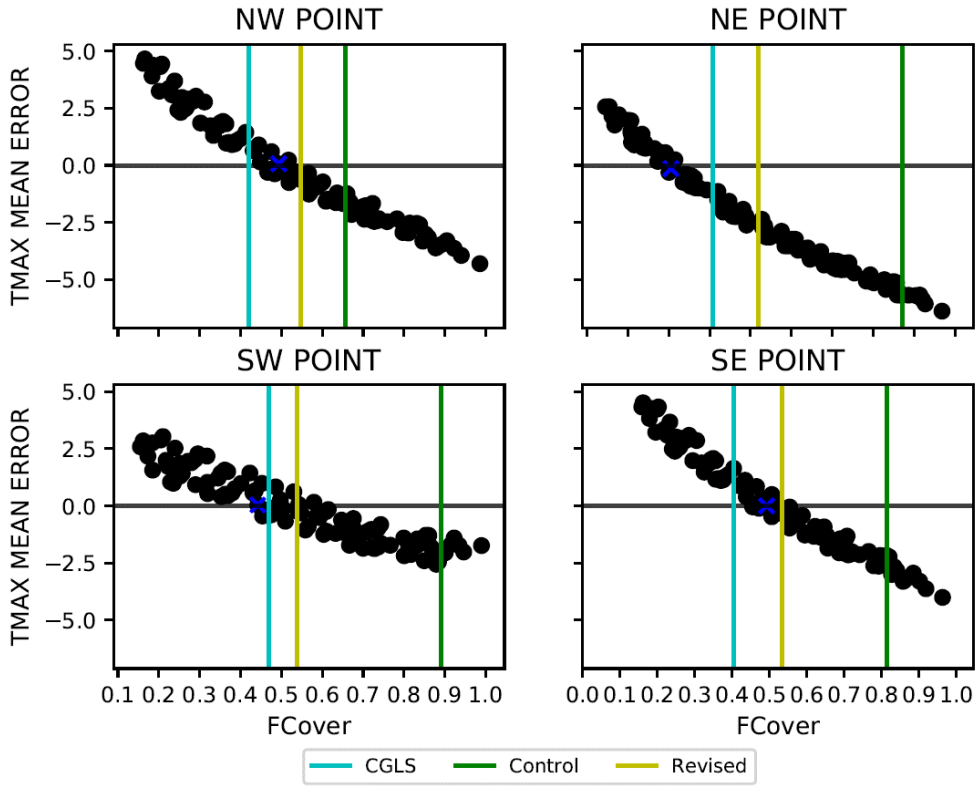


Figure 3.18. Scatterplots of the Tmax Mean Error (y-axis; K) and the FCover (x-axis) of the 100 perturbations (with the revised vegetation) in the Évora-centred domain. The blue cross marks the best perturbation (with Tmax Mean Error closer to zero).

On the other hand, the sensitivity study applied to the c_{veg} parameter for the revised CVL and CVH presents a considerable correlation between the FCover and Tmax Mean Error (Figure 3.18). Except for the NE point, the perturbations with lower Tmax Mean Error are within the range of FCover determined by the CGLS and the Revised FCover. In Figure 3.17, the perturbations with lower errors have, in general, values of low vegetation c_{veg} between 0.4 and 0.6 (except for the NE point), which are similar to the original TVL's c_{veg} of 0.5. The high vegetation c_{veg} can practically take any value because, like CVL in the original vegetation, the CVH in the revised vegetation is quite low.

This study reaffirms the importance of the representation of vegetation in the HTESSEL model, as there is a reasonable correlation between c_{veg} and the Tmax bias with the revised model vegetation, while no correlation is discernible when considering the original model vegetation.

The sensitivity of other model parameters was not assessed due to the shortage of time in the development of this thesis.

4 - Conclusions

The evaluation of the SKT simulation by different ECMWF products (the reanalyses ERA-Interim and ERA5, and the HTESSEL model) in relation to the satellite-LST over the Iberian Peninsula was the first objective of this thesis. In general, the products reasonably underestimate the daytime SKT and slightly overestimate it at night, among the different clusters obtained with the K-Means algorithm, especially in semi-arid regions. In line with previous studies, the results indicate that there is still room for improvement in the simulation of SKT in ECMWF products. Still, ERA5 presents an overall higher quality product in relation to ERA-Interim.

The second objective was to improve the simulation of SKT in the HTESSEL model by running offline simulations in a confined domain with a revised version of certain model parameters. A new representation of vegetation (with new CVL and CVH) based on the ESA-CCI Land Cover dataset provides an overall enhancement in the simulation of SKT during the day. A new soil discretization scheme, on the other hand, just slightly impacts the SKT in a positive way.

A sensitivity study on the c_{veg} parameter provides further proof for a needed revision of the vegetation in the HTESSEL model, as there is a reasonable correlation between c_{veg} and the Tmax mean error with the revised model vegetation (while the same correlation cannot be reproduced with the original model vegetation).

However, the results present a few drawbacks or unsolved issues. Although the new model vegetation leads to improved SKT results, the TVL of the ESA-CCI dataset is not a direct match to the HTESSEL TVL and experimenting with different TVLs translates to poorer results. The phase difference and the nighttime bias observed in the SKT mean diurnal cycle still remain after applying the changes in vegetation and soil discretization.

This study performed uncoupled simulations to assess the impact of surface parameters in the simulation of SKT. It is important to study the effect of the new vegetation in coupled simulations as well. Additionally, due to the satellite-LST relying on IR measurements, the SKT is only assessed in clear sky weather conditions and the conclusions are drawn from that fact. Furthermore, changes in vegetation cover impact the water budget (which was not assessed) and also induce changes in other seasons. However, the reduced satellite-LST availability in the rest of the year and the lack of other observations (e.g. fluxes, soil temperature, soil moisture) limit further diagnostics.

Nevertheless, these results suggest the need to revise the modelled vegetation over the Iberian Peninsula, namely the fraction of low and high vegetation cover in each grid point. Likewise, the definition of the different types of low and high vegetation in the HTESSEL 'lookup' table and the associated c_{veg} values might need a revision as well.

This thesis shows that ERA5 is an improvement over ERA-Interim in simulating SKT. ERA5 still presents considerable SKT daytime biases in relation to satellite data, biases that can be minimised with the implementation of a revised representation of vegetation in the HTESSEL model.

5 - References

Albergel C, E Dutra, S Munier, JC Calvet, JM Sabater, P de Rosnay, and G Balsamo (2018). ERA-5 and ERA-Interim driven ISBA land surface model simulations: which one performs better?, *Hydrol. Earth. Syst. Sci.*, 22, 3515-3532, doi:10.5194/hess-22-3515-2018.

Balsamo G, S Boussetta, E Dutra, A Beljaars, P Viterbo, and BV der Hurk. Evolution of land-surface processes in the IFS, *ECMWF newsletter*, 127, 17-22, 2011, doi:10.21957/x1j3i7bz.

Balsamo G, C Albergel, A Beljaars, S Boussetta, E Brun, H Cloke, D Dee, E Dutra, J Muñoz-Sabater, F Pappenberger, P de Rosnay, T Stockdale, and F Vitart (2015). ERA-Interim/Land: a global land surface reanalysis data set, *Hydrol. Earth Syst. Sci.*, 19, 389–407, doi:10.5194/hess-19-389-2015.

Belo-Pereira M, E Dutra, and P Viterbo (2011). Evaluation of global precipitation data sets over the Iberian Peninsula, *J. Geophys. Res.*, 116, D20101, doi:10.1029/2010JD015481.

Best MJ, G Abramowitz, HR Johnson, AJ Pitman, G Balsamo, A Boone, M Cuntz, B Decharme, PA Dirmeyer, J Dong, M Ek, Z Guo, V Haverd, BJJ van den Hurk, GS Nearing, B Pak, C Peters-Lidard, JA Santanello Jr., L Stevens, and N Vuichard (2015). The Plumbing of Land Surface Models: Benchmarking Model Performance, *J. Hydrometeorol.*, 16, 1425–1442, doi:10.1175/JHM-D-14-0158.1.

Bojinski S, M Verstraete, TC Peterson, C Richter, A Simmons, and M Zemp (2014). The concept of essential climate variables in support of climate research, applications, and policy, *Bull. Amer. Meteor. Soc.*, 95, 1431-1443, doi:10.1175/BAMS-D-13-00047.1.

Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), last access: 18 February 2019. https://cds.climate.copernicus.eu/cdsapp#!/home

Dee DP, SM Uppala, AJ Simmons, P Berrisford, P Poli, S Kobayashi, U Andrae, MA Balmaseda, G Balsamo, P Bauer, P Bechtold, ACM Beljaars, L van de Berg, J Bidlot, N Bormann, C Delsol, R Dragani, M Fuentes, AJ Geer, L Haimberger, SB Healy, H Hersbach, EV Hólm, L Isaksen, P Kållberg, M Köhler, M Matricardi, AP McNally, BM Monge-Sanz, J-J Morcrette, B-K Park, C Peubey, P de Rosnay, C Tavolato, J-N Thépaut, and F Vitart (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quart. J. Roy. Meteor. Soc.*, 137, 553–597, doi:10.1002/qj.828.

Ghent D, J Kaduk, J Remedios, J Ardö, and H Balzter (2010). Assimilation of land surface temperature into the land surface model JULES with an ensemble Kalman filter, *J. Geophys. Res.*, 115, D19112, doi:10.1029/2010JD014392.

Good EJ (2016). An in situ-based analysis of the relationship between land surface "skin" and screen-level air temperatures, *J. Geophys. Res. Atmos.*, 121, 8801-8819, doi:10.1002/2016JD025318.

Han X, S Duan, C Huang, and Z Li (2018). Cloudy land surface temperature retrieval from three-channel microwave data, *International Journal of Remote Sensing*, 40:5-6, 1793-1807, doi:10.1080/01431161.2018.1471552.

Hersbach H, P de Rosnay, B Bell, D Schepers, A Simmons, C Soci, S Abdalla, M Alonso Balmaseda, G Balsamo, P Bechtold, P Berrisford, J Bidlot, E de Boisséson, M Bonavita, P Browne, R Buizza, P Dahlgren, D Dee, R Dragani, M Diamantakis, J Flemming, R Forbes, A Geer, T Haiden, E Hólm, L Haimberger, R Hogan, A Horányi, M Janisková, P Laloyaux, P Lopez, J Muñoz-Sabater, C Peubey,

R Radu, D Richardson, JN Thépaut, F Vitart, X Yang, E Zsótér and H Zuo (2018). Operational global reanalysis: progress, future directions and synergies with NWP, *ECMWF ERA Report*, 27, doi:10.21957/tkic6g3wm.

Hersbach H, B Bell, P Berrisford, A Horányi, JM Sabater, J Nicolas, R Radu, D Schepers, A Simmons, C Soci, and D Dee. Global reanalysis: goodbye ERA-Interim, hello ERA5, *ECMWF newsletter*, 159, 17-24, 2019, doi:10.21957/vf291hehd7.

Kala J, MG De Kauwe, AJ Pitman, BE Medlyn, YP Wang, R Lorenz, and SE Perkins-Kirkpatrick (2016). Impact of the representation of stomatal conductance on model projections of heatwave intensity, *Sci. Rep.*, 6, 23418, doi:10.1038/srep23418.

Loveland TR, BC Reed, JF Brown, DO Ohlen, Z Zhu, L Youing, and JW Merchant (2000). Development of a global land cover characteristics database and IGB6 DISCover from the 1km AVHRR data, *Int. J. Remote Sensing*, 21, 1303–1330, doi:10.1080/014311600210191.

Norman JM, and F Becker (1995). Terminology in thermal infrared remote sensing of natural surfaces, *Agricultural and Forest Meteorology*, 77, 153–166, doi:10.1016/0168-1923(95)02259-Z

Orth R, E Dutra, IF Trigo, and G Balsamo (2017). Advancing land surface model development with satellite-based Earth observations, *Hydrol. Earth Syst. Sci.*, 21, 2483-2495, doi:10.5194/hess-21-2483-2017.

Poulter B, N MacBean, A Hartley, I Khlystova, O Arino, R Betts, S Bontemps, M Boettcher, C Brockman, P Defourny, S Hagemann, M Herold, G Kirches, C Lamarche, D Lederer, C Ottlé, M Peters, and P Peylin (2015). Plant functional type classification for earth system models: results from the European Space Agency's Land Cover Climate Change Initiative, *Geosci. Model Dev.*, 8, 2315-2328, doi:10.5194/gmd-8-2315-2015.

Prigent C, C Jimenez, and F Aires (2016). Toward "all weather", long record, and real-time land surface temperature retrievals from microwave satellite observations, *J. Geophys. Res. Atmos.*, 121, 5699-5717, doi:10.1002/2015JD024402.

Schär C, PL Vidale, D Lüthi, C Frei, C Häberli, MA Liniger, and C Appenzeller (2004). The role of increasing temperature variability in European summer heatwaves, *Nature*, 427, 332-336, doi:10.1038/nature02300.

Seneviratne SI, D Lüthi, M Litschi, and C Schär (2006). Land-atmosphere coupling and climate change in Europe, *Nature*, 443, 205-209, doi:10.1038/nature05095.

Sobol IM (1967). On the distribution of points in a cube and the approximate evaluation of integrals, *USSR Comput. Math. Math.*, 7, 86–112, doi:10.1016/0041-5553(67)90144-9.

Trigo IF, CC Dacamara, P Viterbo, JL Roujean, F Olesen, C Barroso, F Camacho-de-Coca, D Carrer, SC Freitas, J García-Haro, B Geiger, F Gellens-Meulenberghs, N Ghilain, J Meliá, L Pessanha, N Siljamo, and A Arboleda (2011). The Satellite Application Facility for Land Surface Analysis, *International Journal of Remote Sensing*, 32:10, 2725-2744, doi:10.1080/01431161003743199.

Trigo IF, S Bousseta, P Viterbo, G Balsamo, A Beljaars, and I Sandu (2015). Comparison of model land skin temperature with remotely sensed estimates and assessment of surface-atmosphere coupling, *J. Geophys. Res. Atmos.*, 120, 12,096-12,111, doi:10.1002/2015JD023812.

Ukkola AM, MG De Kauwe, AJ Pitman, MJ Best, G Abramowitz, V Haverd, M Decker, and N Haughton (2016). Land surface models systematically overestimate the intensity, duration and

magnitude of seasonal-scale evaporative droughts, *Environ. Res. Lett.*, 11, 104012, doi:10.1088/1748-9326/11/10/104012.

Verger A, F Baret, and M Weiss (2014). Near real-time vegetation monitoring at global scale, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7, 3473-3481, doi:10.1109/JSTARS.2014.2328632.

Viterbo P, and A Beljaars (1995). An improved land surface parameterization scheme in the ECMWF model and its validation, *J. Climate*, 8, 2716-2748, doi:10.1175/1520-0442(1995)008<2716:AILSPS>2.0.CO;2.

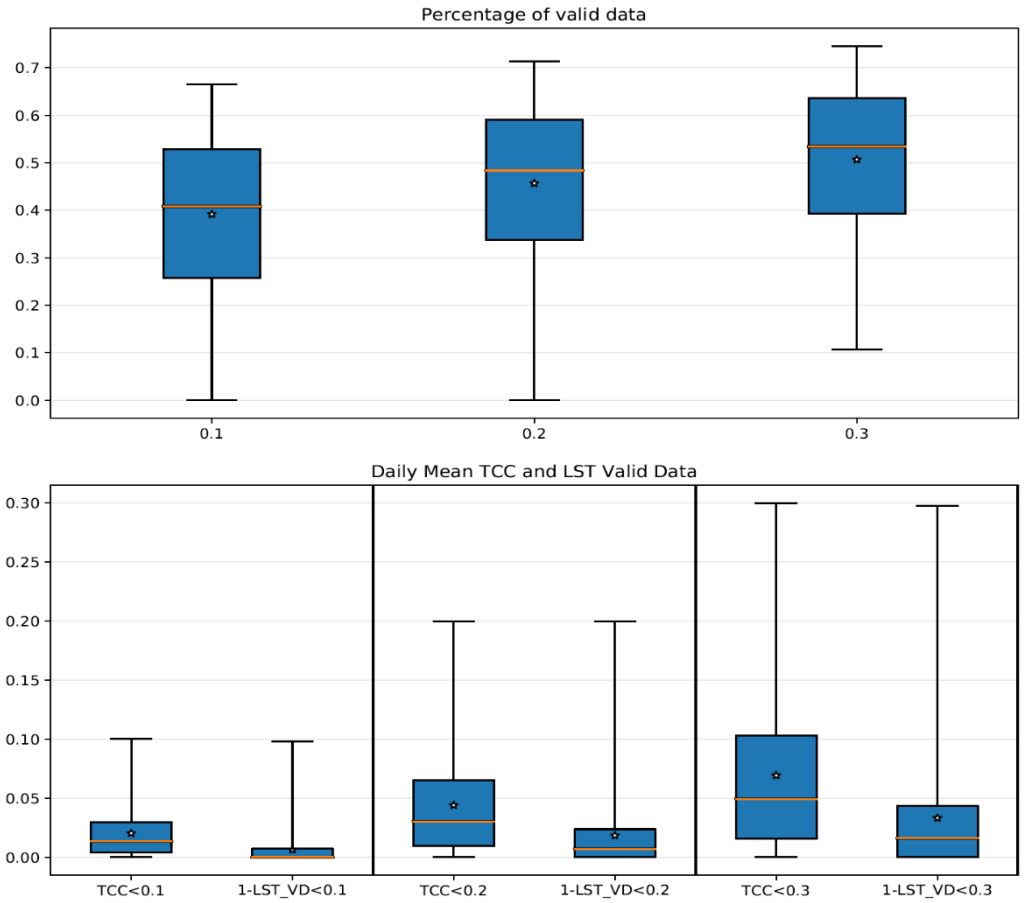
Wan Z (1999). MODIS Land-Surface Temperature Algorithm Theoretical Basis Document In. Greenbelt MD, USA: NASA/GSFC.

Zhou C, K Wang, and Q Ma (2017). Evaluation of Eight Current Reanalyses in Simulating Land Surface Temperature from 1979 to 2003 in China, *J. Climate*, 30, 7379-7398, doi:10.1175/JCLI-D-16-0903.1.

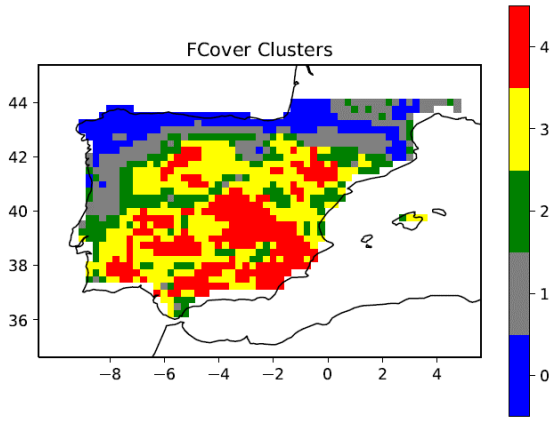
Table 2. Default land cover to plant functional type cross-walking table provided by the conversion tool with the 22 level 1 UNLCCS classes and 14 level 2 UNLCCS sub-classes in italics. The units are % coverage of each PFT per UNLCCS class.

LCCS	UNLCCS Land Cover Class Description	Tree				Shrub				Grass		Non-vegetated		
Class		BrEv	BrDc	NeEv	NeDe	BrEv	BrDc	NeEv	NeDe	Nat. Grass	Man. Grass	Bare soil	Water	Snow/Ice
10	Cropland, rainfed										100			
11	Herbaceous cover										100			
12	Tree or shrub cover						50				50			
20	Cropland, irrigated or post-flooding										100			
30	Mosaic cropland (> 50 %) nat. veg. (tree, shrub, herb.) (< 50 %)	5	5			5	5	5		15	60			
40	Mosaic nat. veg. (tree, shrub, herb.) (> 50 %)/cropland (< 50 %)	5	5			7.5	10	7.5		25	40			
50	Tree cover, broadleaf, evergreen, closed to open (> 15 %)	90				5	5							
60	Tree cover, broadleaf, deciduous, closed to open (> 15 %)		70				15			15				
61	Tree cover, broadleaf, deciduous, closed (> 40 %)		70				15			15				
62	Tree cover, broadleaf, deciduous, open (15–40 %)		30				25			35		10		
70	Tree cover, needleleaf, evergreen, closed to open (> 15 %)			70		5	5	5		15				
71	Tree cover, needleleaf, evergreen, closed (> 40 %)			70		5	5	5		15				
72	Tree cover, needleleaf, evergreen, open (15–40 %)			30			5	5		30		30		
80	Tree cover, needleleaf, deciduous, closed to open (> 15 %)				70	5	5	5		15				
81	Tree cover, needleleaf, deciduous, closed (> 40 %)				70	5	5	5		15				
82	Tree cover, needleleaf, deciduous, open (15–40 %)				30		5	5		30		30		
90	Tree cover, mixed leaf type (broadleaf and needleleaf)		30	20	10	5	5	5		15		10		
100	Mosaic tree and shrub (> 50 %)/herbaceous cover (< 50 %)	10	20	5	5	5	10	5		40				
110	Mosaic herbaceous cover (> 50 %)/tree and shrub (< 50 %)	5	10	5		5	10	5		60				
120	Shrubland					20	20	20		20		20		
121	Shrubland evergreen					30		30		20		20		
122	Shrubland deciduous						60			20		20		
130	Grassland									60		40		
140	Lichens and mosses									60		40		
150	Sparse vegetation (tree, shrub, herbaceous cover) (< 15 %)	1	3	1		1	3	1		5		85		
152	Sparse shrub (< 15 %)					2	6	2		5		85		
153	Sparse herbaceous cover (< 15 %)									15		85		
160	Tree cover, flooded, fresh or brackish water	30	30							20			20	
170	Tree cover, flooded, saline water	60				20							20	
180	Shrub/herbaceous cover, flooded, fresh/saline/brackish water		5	10			10	5		40			30	
190	Urban areas		2.5	2.5						15		75	5	
200	Bare areas											100		
201	Consolidated bare areas											100		
202	Unconsolidated bare areas											100		
210	Water bodies												100	
220	Permanent snow and ice													100

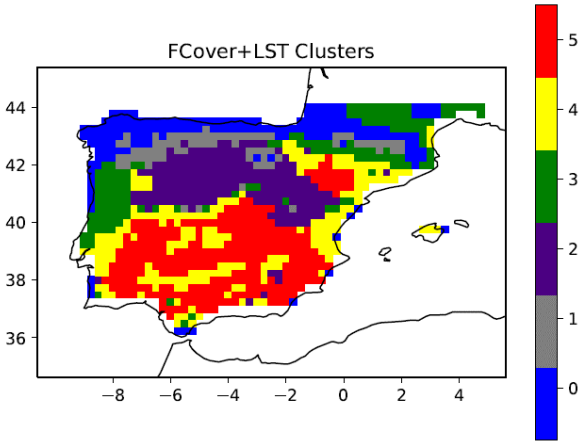
Appendix 6.1. Cross-walking table suggested by Poulter et al. (2015).



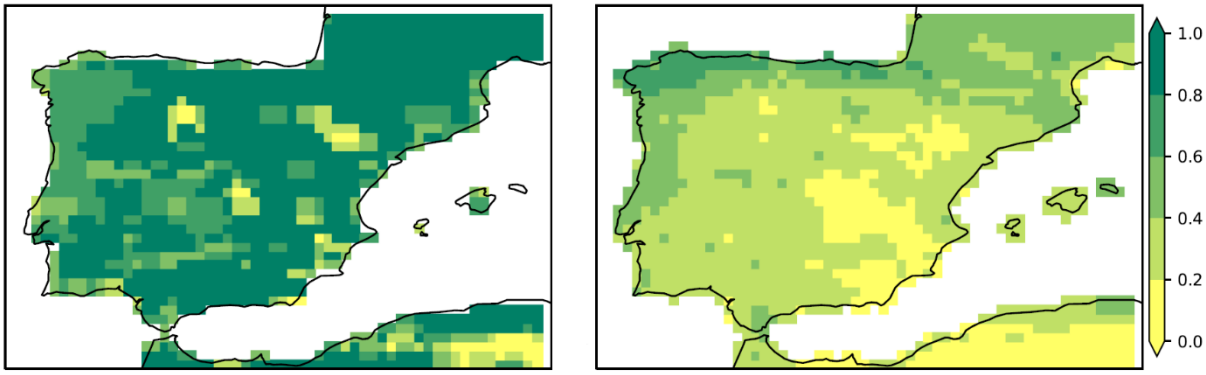
Appendix 6.2. Mean percentage of data that follows the double threshold (top) and the Daily Mean TCC and LST valid data (LST_VD) (bottom) in each grid point, according to three thresholds: 0.1 (left), 0.2 (center) and 0.3 (right).



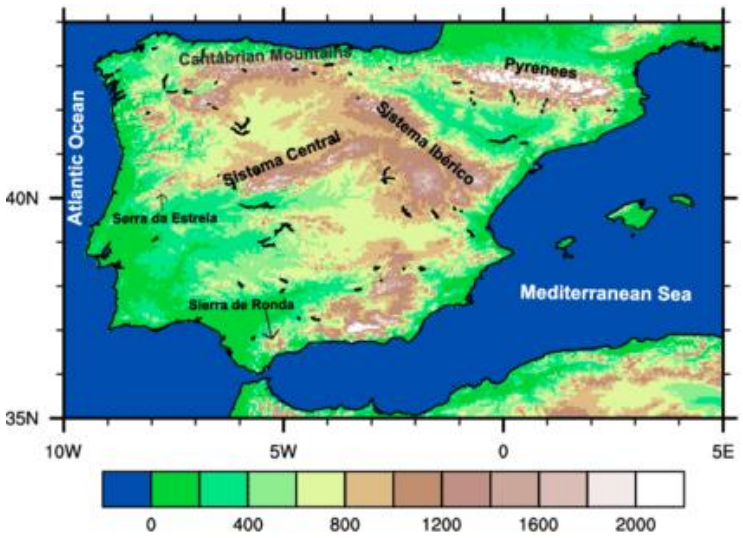
Appendix 6.3. Clusters determined by the K-Means Algorithm, with CGLS-FCover as input.



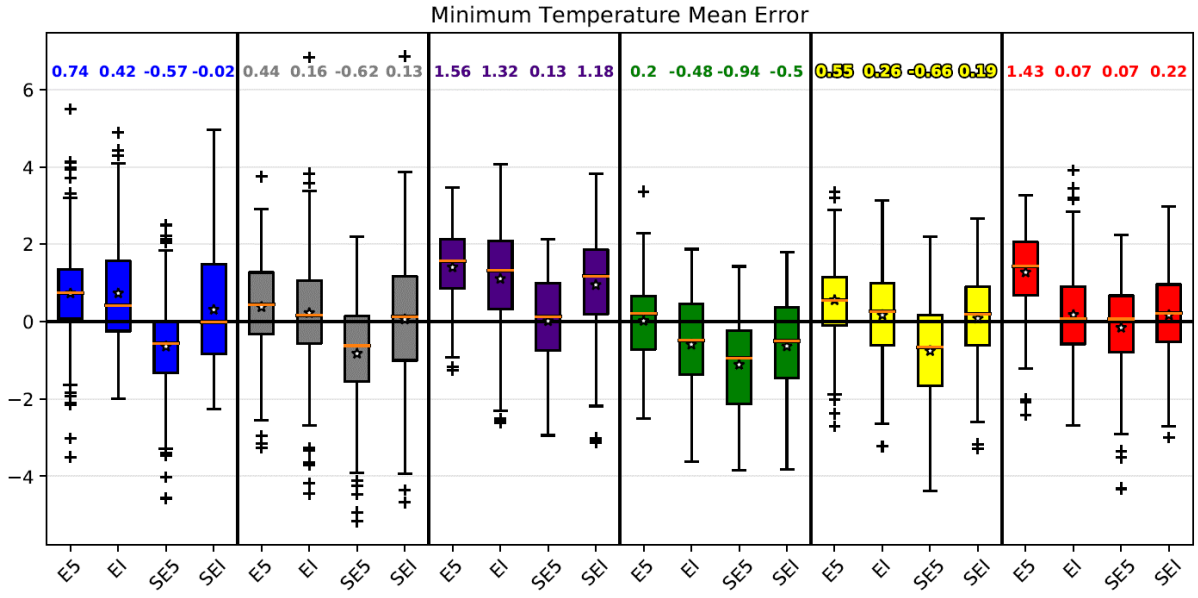
Appendix 6.4. Clusters determined by the K-Means Algorithm, with both LST and CGLS-FCover as input.



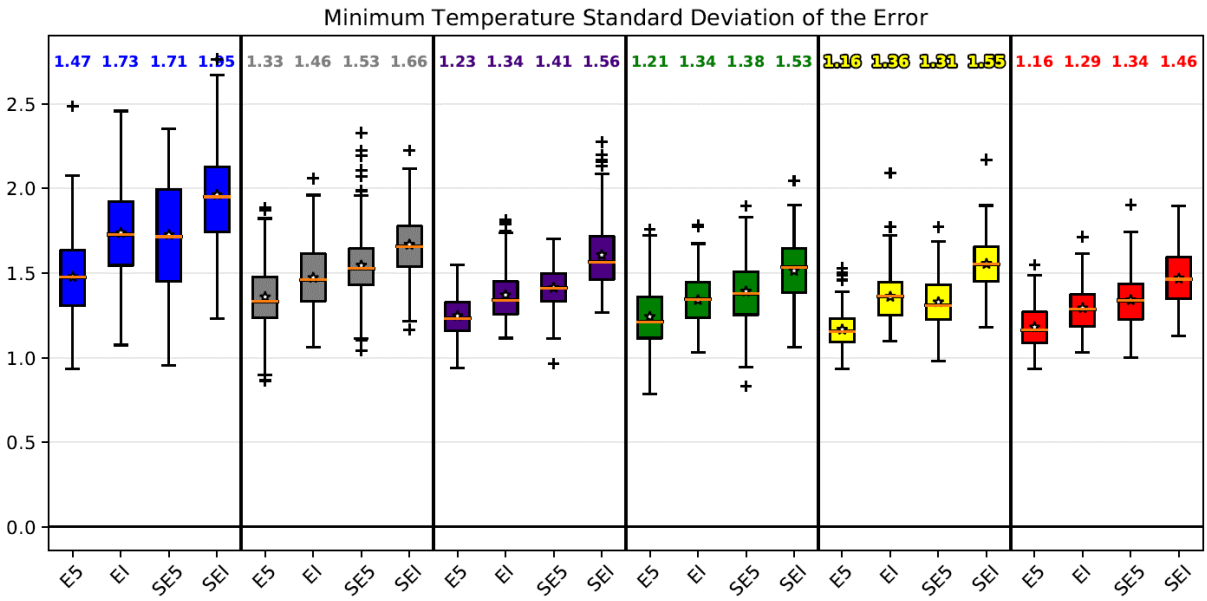
Appendix 6.5. HTESSEL FCover (left) and CGLS 1999-2018 Mean FCover (right) in Iberia.



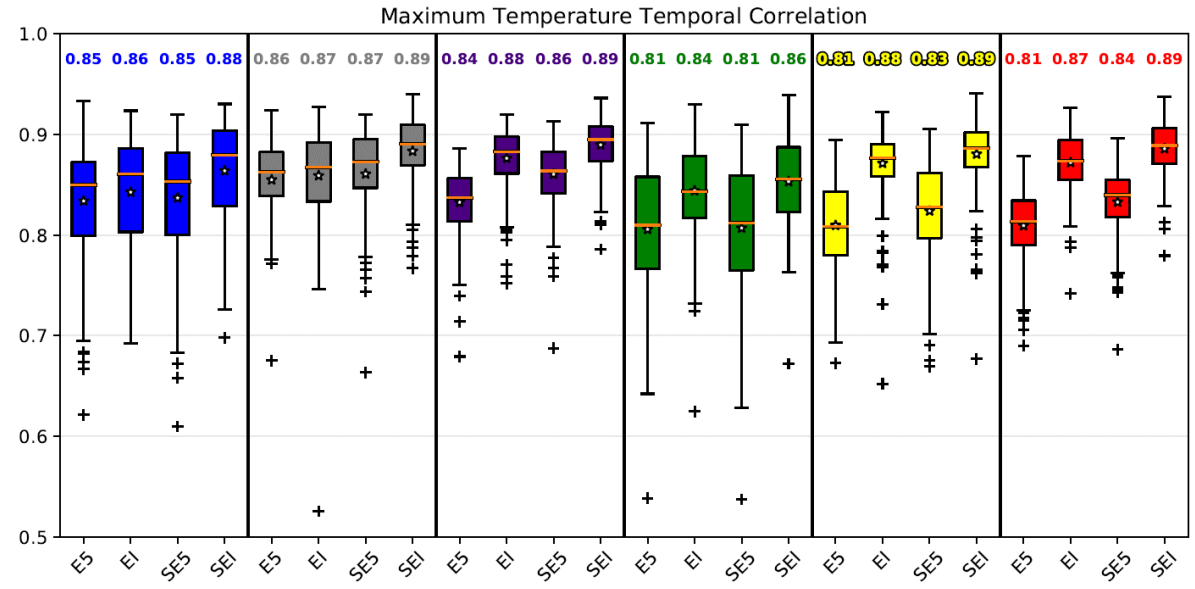
Appendix 6.6. Main orographic features of the Iberian Peninsula (from Belo-Pereira et al., 2011). Altitude in metres.



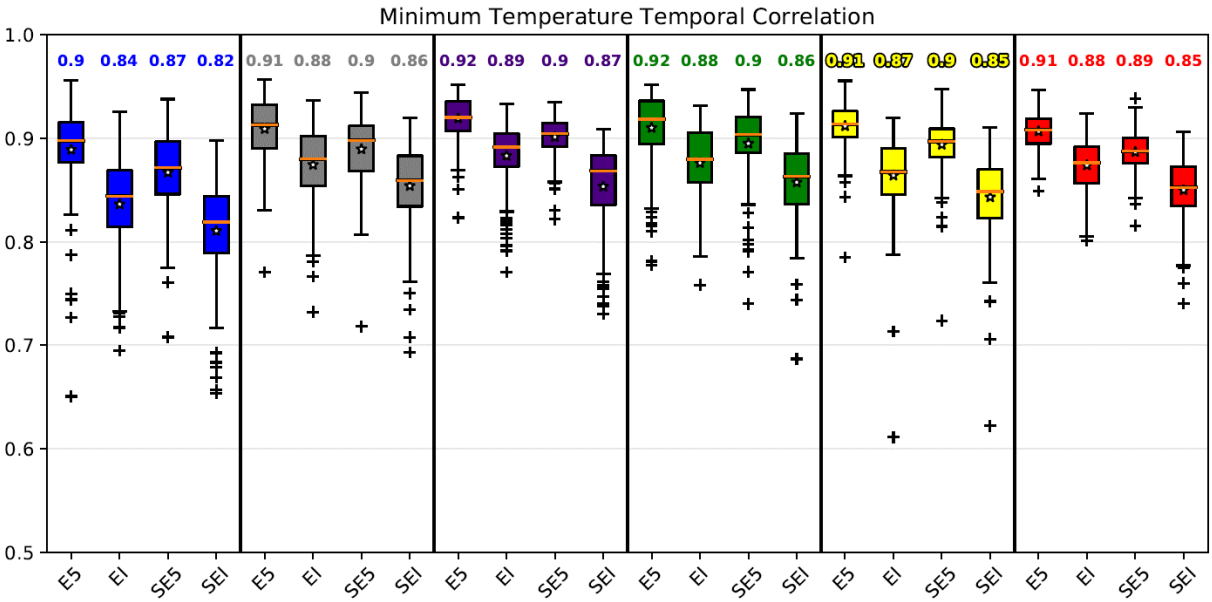
Appendix 6.7. Tmin Bias (K) in: ERA5 (E5), ERA-Interim (EI) and the simulations forced with E5 (SE5) and EI (SEI). The colours represent the clusters in Figure 3.1. The number above each boxplot is the median of that boxplot.



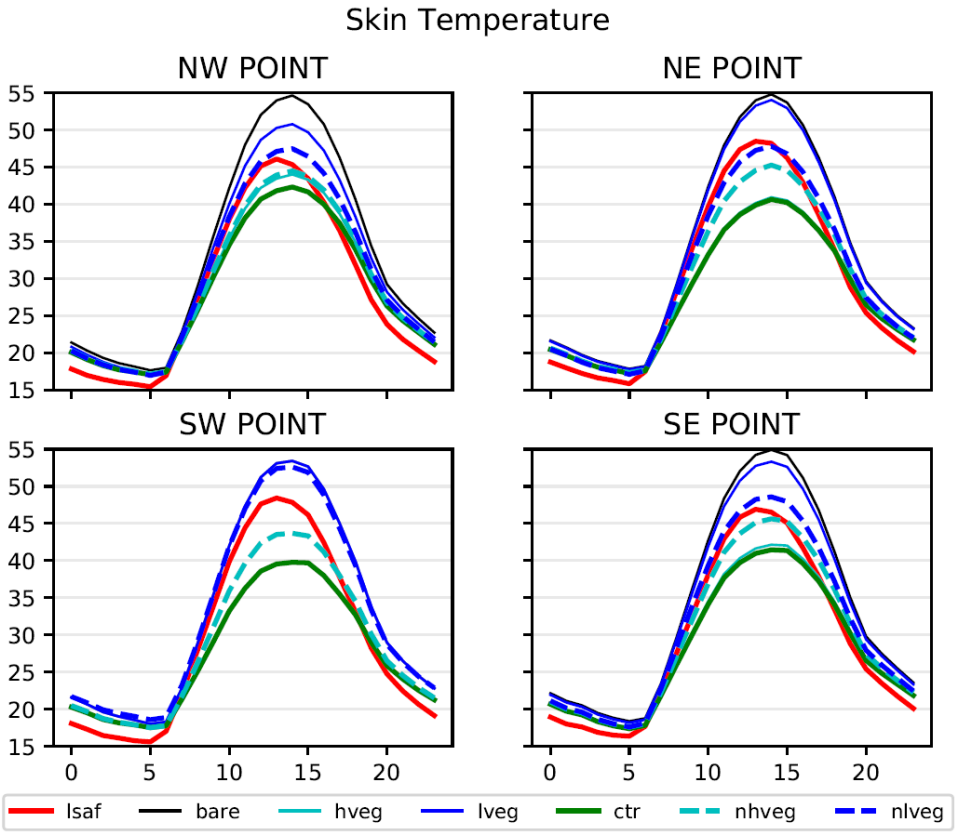
Appendix 6.8. Tmin SD (K) in: ERA5 (E5), ERA-Interim (EI) and the simulations forced with E5 (SE5) and EI (SEI). The colours represent the clusters in Figure 3.1. The number above each boxplot is the median of that boxplot.



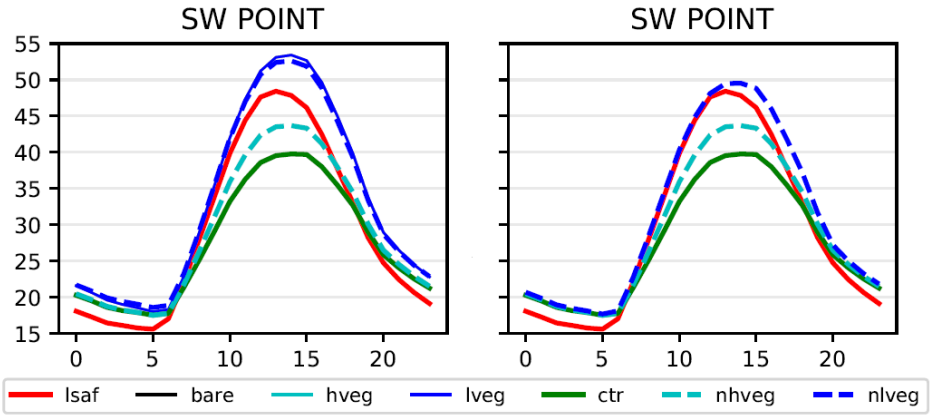
Appendix 6.9. Tmax Correlation in: ERA5 (E5), ERA-Interim (EI) and the simulations forced with E5 (SE5) and EI (SEI). The colours represent the clusters in Figure 3.1. The number above each boxplot is the median of that boxplot.



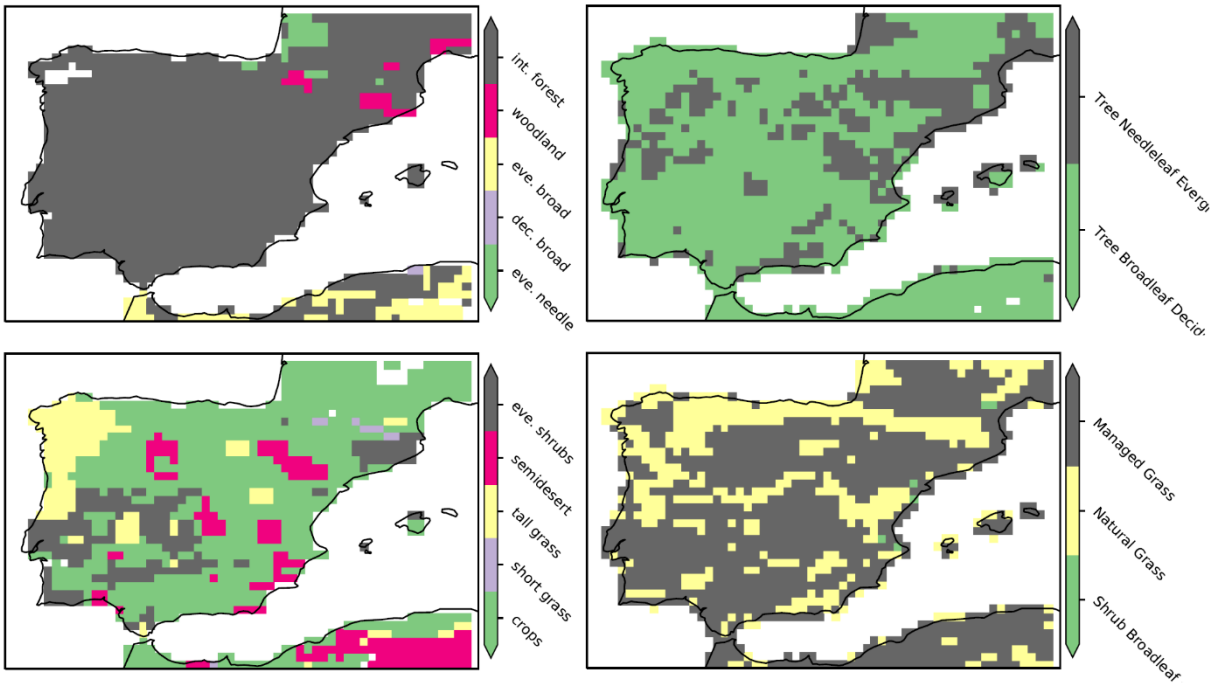
Appendix 6.10. Tmin Correlation in: ERA5 (E5), ERA-Interim (EI) and the simulations forced with E5 (SE5) and EI (SEI). The colours represent the clusters in Figure 3.1. The number above each boxplot is the median of that boxplot.



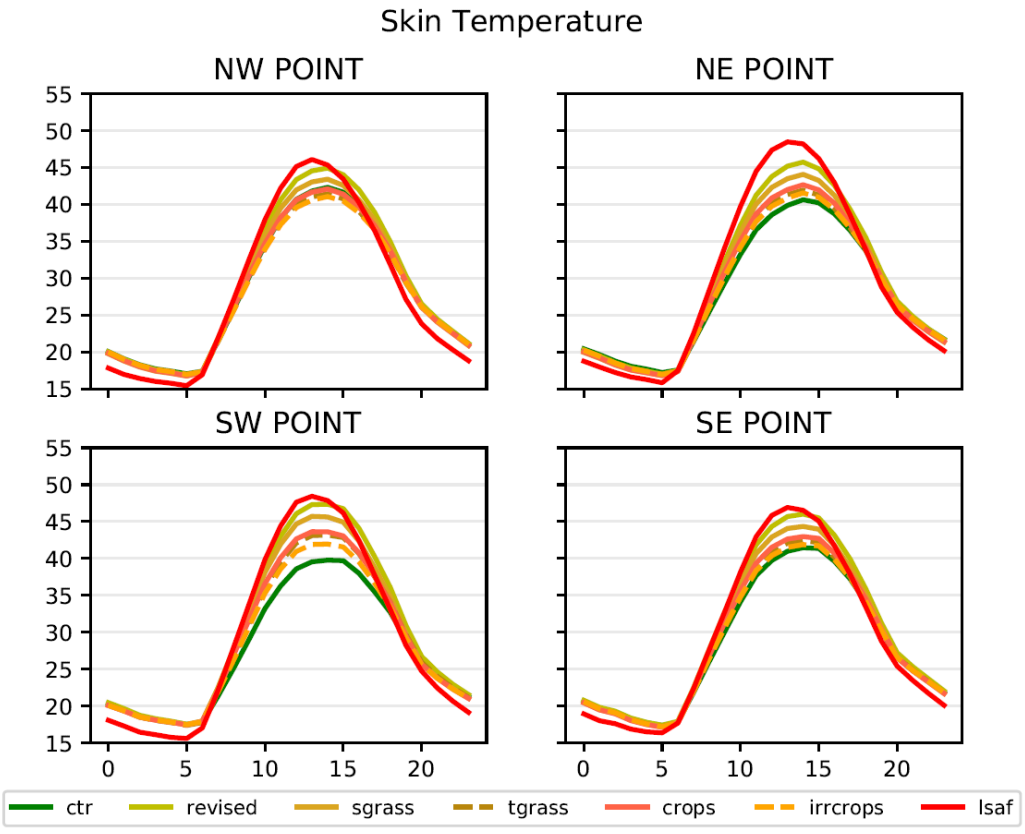
Appendix 6.11. 2010 JJA Mean SKT Diurnal Cycle (°C) in the Évora-centred domain: 'bare' (black), 'hveg' (light blue, solid line), 'lveg' (dark blue, solid line), 'nhveg' (light blue, dashed line) and 'nlveg' (dark blue, dashed line) simulations.



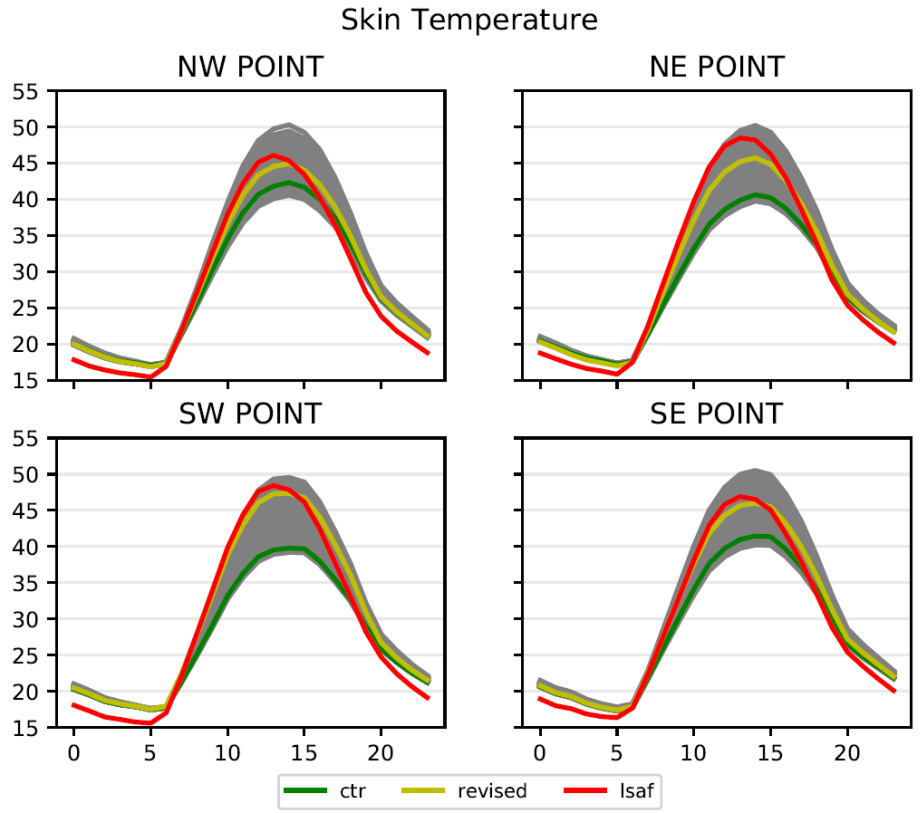
Appendix 6.12. 2010 JJA Mean SKT Diurnal Cycle (°C) in the SW point of the Évora-centred domain, 'nlveg' with original TVL (left) and revised TVL (right).



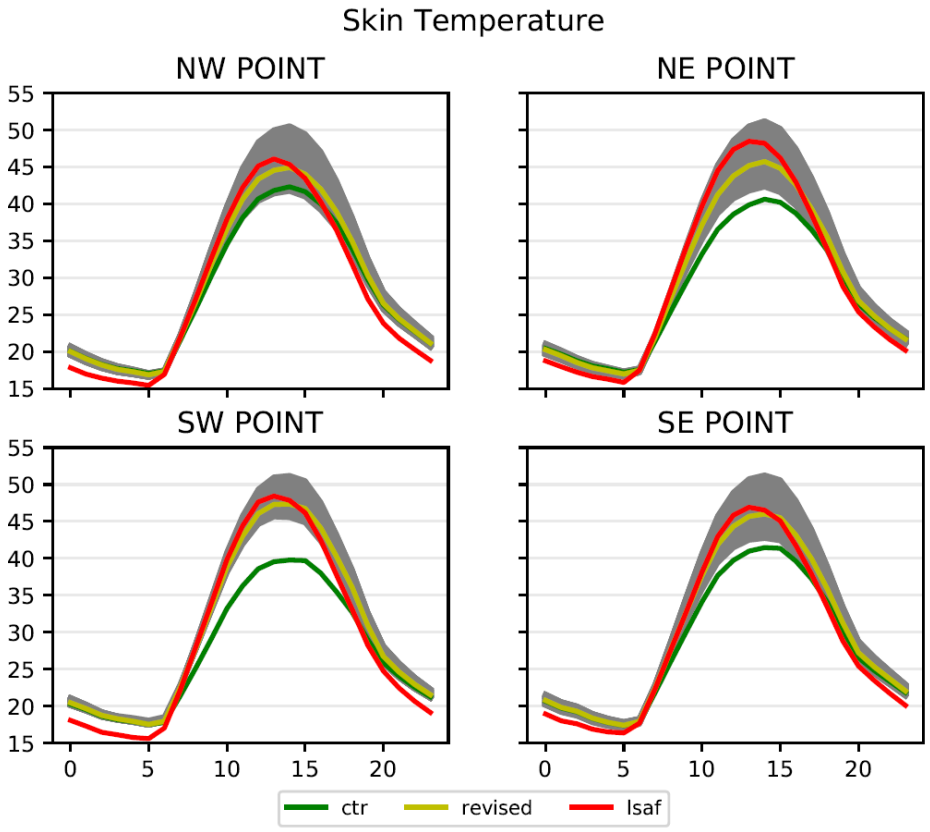
Appendix 6.13. HTESSEL's (left) and ESA-CCI's (as PFTs) (right) TVH (top) and TVL (bottom).



Appendix 6.14. 2010 JJA Diurnal Cycle of SKT (°C): TVL experiments - short grass (dark yellow), tall grass (brown), crops (dark orange) and irrigated crops (light orange). Also represented are the control (green) and revised (yellow) simulations and satellite-LST (red).



Appendix 6.15. 2010 JJA Mean SKT Diurnal Cycle (°C) in the Évora-centred domain: satellite-lst (red), control (green) and revised (yellow) simulations, and the 100 perturbations with the original vegetation (grey).



Appendix 6.16. 2010 JJA Mean SKT Diurnal Cycle (°C) in the Évora-centred domain: satellite-lst (red), control (green) and revised (yellow) simulations, and the 100 perturbations with the revised vegetation (grey).